Winter 2008

Evaluating accuracy issues in mapping benthic habitats: An investigation in the causes of misclassification and the importance of segmentation parameters

Meghan E. Graham
University of New Hampshire, Durham

EVALUATING ACCURACY ISSUES IN MAPPING BENTHIC HABITATS:
AN INVESTIGATION IN THE CAUSES OF MISCLASSIFICATION AND THE
IMPORTANCE OF SEGMENTATION PARAMETERS

BY

MEGHAN E. GRAHAM

B.S., Clarkson University, 2006

THESIS

Submitted to the University of New Hampshire
in Partial Fulfillment of
the Requirements for the Degree of

Master of Science
in
Natural Resources

December, 2008

This thesis has been examined and approved.

Thesis Director, Dr. Russell G. Congalton
Professor of Remote Sensing and Geographic Information Systems

Dr. Mark J. Ducey, Professor of Forest Biometrics and Management

Ms. Kass Green, Research Scientist and President of the Alta Vista Company

December 10, 2008

Date

ACKNOWLEDGEMENTS

I would like to sincerely thank my advisor, Dr. Russ Congalton, for all of his help and guidance throughout this project. His belief in me, as well as his support, has made this a very wonderful and informative experience. I would also like to thank my committee members: Dr. Mark Ducey, for his unwavering kindness and statistical expertise; and Ms. Kass Green, for introducing me to this project and providing me with any information I could possibly need throughout this research endeavor. I would also like to extend a thank you to Chad Lopez of Furgo EarthData for sending me and helping me understand this data, as well as NOAA for allowing me to work on this project.

I am also very grateful for all the staff and students at UNH who have supported me throughout my time here. They have made this work fun and painless. I would also like to thank all of the students of BASAL, especially Brianna Heath, for spending hours with me bouncing around ideas.

Finally, I cannot properly express my how much I am indebted to my friends and my family for supporting me throughout my life and helping me finish my masters. A special thank you goes to my mom for reading my thesis almost as many times as I have. Also, thank you Rich for making me breakfast, lunch, and dinner in these final few months, I am not sure I would have eaten without you.

Thank you to the MS-60 NH Agriculture Experiment Station Grant (Congalton) and the 2008 Summer TA Fellowship given to me by the UNH Graduate School for providing me with the means to make this research possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

EVALUATING ACCURACY ISSUES IN MAPPING BENTHIC HABITATS: AN INVESTIGATION IN THE CAUSES OF MISCLASSIFICATION AND THE IMPORTANCE OF SEGMENTATION PARAMETERS

By

Meghan E. Graham

University of New Hampshire, December, 2008

Benthic habitats are some of the most difficult habitats to map using remote sensing. In a study of six bays in the Texas Gulf Coast, maps were created from digital aerial imagery. Using an object-based image analysis (OBIA) approach, the image was classified with the Classification and Regression Tree (CART) technique to produce a draft map. The draft map was then extensively manually edited to produce a contractor map. Accuracy assessments of both maps revealed that the two were not significantly different. The objective of this study was to determine why the editing did not improve the draft map. Our analyses indicate that the small segmentation scale parameter chosen for the map over-segmented the image and reduced the effectiveness of the classification technique and the manual editing. When compared to a similar map with a larger scale parameter, the smaller initial polygons proved more difficult to accurately classify.

CHAPTER I

INTRODUCTION

The fragility of coastal marine or benthic habitats all over the world is of growing concern to both scientists and coastal communities alike. In Texas, these habitats are particularly important because they support a vast variety of marine ecosystems, as well as provide a number of recreational opportunities and contribute considerably to the Texas economy (Green and Lopez, 2007). In the Gulf of Mexico, seagrass is a critical element of the marine ecosystems for many reasons. Seagrass is a vital component of nursery habitats because it provides food and shelter for juvenile fish, turtles, and invertebrates. Seagrass also contributes to the surrounding ecosystem by improving water clarity and quality by acting as a filtration system. This seemingly trivial vegetation also provides a very important buffer to the shoreline by dispelling wave energy that could cause significant erosion and destruction of the coastal area (Koch, 2001; Handley et al., 2007). However, the 2-3% growth of the coastal population per year in Texas has impacted the bay ecosystems in such a way as to result in the loss of these habitats (Green et al., 2007).

Over the past few years, a group of scientists has begun working to identify and classify the benthic habitats in bay areas along the coast of Texas based on the Florida System for Classification of Habitats in Estuarine and Marine Environments (SCHEME)

(Madley et al., 2002). Phase I of the project was completed in July of 2007 and involved the classification of the habitats in six different coastal areas, encompassing about 1,400 square miles of the Gulf of Mexico coast (Green et al., 2007). Phase II of the project includes two remaining large bays on the southern coast of Texas and is expected to be complete by the end of 2008 (Green et al., 2007). The goal of this project was to generate baseline benthic maps of these areas to begin to monitor changes in seagrasses and other vital components of the bay ecosystems. These maps, when compared to future maps, could then be used to measure the gains and losses in the Gulf of Mexico of benthic habitats over time, and focus conservation efforts on areas of need. Since these maps are to be used quantitatively year after year, it is extremely important that the maps be reproducible for future studies and be highly accurate (Congalton and Green, 2009).

One of the most efficient ways to map areas of this magnitude is to use aerial or satellite imagery of the study site. Generally, there are three different ways of creating a map from photos or images. The first and oldest method for creating a map is through manual photo-interpretation, where a photo-interpreter uses his or her judgment to delineate and classify the habitats in an image. This method can either be done by drawing habitat outlines by hand on paper, or on the computer using a Geographic Information System (GIS) (Jensen, 2005). The second method is through pixel-based classification, where each pixel on an image is classified individually and separately from its surrounding pixels (Blaschke and Strobl, 2001; Jensen, 2005). The third method is an object-based image analysis (OBIA) approach, where individual pixels with similar properties are first grouped into polygons and then these polygons are individually classified. The action of grouping similar pixels is usually referred to as segmenting the image. Typically, the intent is to segment the image so that pixels that are close to each other and in the same habitat are grouped together and do not include pixels of another habitat nearby (Baatz et al., 2001; Jensen, 2005). These last two methods, the pixel-based and OBIA approaches, are usually implemented using a variety of computer-based classification algorithms. These algorithms fall into three general categories: supervised classification algorithms; unsupervised classification algorithms; and a hybrid of the first two classification algorithms (Jensen, 2005; Congalton and Green, 2009).

In order to create precise benthic habitat maps for the Phase I study area, digital aerial imagery was collected of the area at 2 meter pixel resolution. The researchers that completed Phase I of this project then chose to use an OBIA approach to classify this imagery and segmented the images into very small initial polygons. Each of the polygons was then classified using a classification algorithm called a Classification and Regression Tree (CART) technique. The CART technique uses data collected by the research team, either through site visits or photo-interpretation, to create a decision tree that automatically classifies each polygon on the imagery. This approach to creating a benthic habitat map is less expensive and time consuming since it does not involve hours of manual photo-interpretation (Green and Lopez, 2007).

With minimal human involvement and correction, the initial draft benthic habitat map for Phase I had a deterministic accuracy of 83% (Green et al., 2007). The draft map was then extensively manually edited in hopes of increasing the accuracy of the benthic habitat map. During the manual editing process, many polygons were changed from one label to another due to human knowledge of the area and results from photo-interpretation. Small polygons were also dissolved into larger polygons with the same label. Overall, around 26% of the polygons changed labels during the 500 or more person-hours of editing. The result of this editing was a new 'contractor' map. Although the labels of many of the polygons were quite different, the new deterministic accuracy of the contractor map was only 85%, which is not significantly better than the original draft map (Green et al., 2007).

The main objective for my study was to determine what could have caused this difference in labeling of polygons without a change in accuracy. Specifically, this paper examines three different alternative and null hypotheses regarding this occurrence:

H_1: Some of the benthic habitat classes were spectrally and/or visually inseparable on the imagery causing a large change in polygon labels of certain habitats during editing.

$H_0 1$: No one group of changed polygons is a different size than any other group of changed polygons.

H_2: The accuracy sites were statistically more often in areas of homogeneous habitat causing an artificial inflation of the results of the accuracy assessment.

$H_0 2$: The proportion of accuracy sites within each transition zone is the same as the proportion of area covered by that transition zone.

H_3: The small initial in polygon size in Phase I over-segmented the image making the classification process less effective in comparison to the Phase II maps which were generated with larger initial polygons.

$H_0 3$: The effectiveness of the mapping methods did not change with the size of the segmentation scale parameter.

In order to test these hypotheses, the specific objectives of the study were to:

- Quantify the number of changed polygons between the draft and contractor maps and compute the number of changed polygons for each habitat class.
- Determine if one habitat was changed to another habitat significantly more than any other habitat change during editing.
- Define and label areas where one habitat type is expected to be transitioning into another habitat type and therefore not likely an area of homogeneous habitat.
- Find the portion of changed polygons within each transition zone and determine if a significant fraction of the changed polygons is within the transition zone.
- Compute the percentage of accuracy sites collected in each of the transition zones and compare it to the portion of the total area covered by the transition zone to determine whether there are an adequate number of accuracy sites within each transition zone.
- Compute the average size of polygons changed during editing and compare it with the overall average polygon size of the draft and contractor maps.
- Compare the average sizes of the smaller initial polygons in the Phase I maps with the average sizes of the larger initial polygons in Phase II maps.
- Determine whether there is a difference in the accuracies of the maps created in Phase I and Phase II.
- Assess how the difference in initial polygon size may have affected the creation and accuracy of the final benthic habitat maps in Phase I and Phase II.

CHAPTER II

LITERATURE REVIEW

Background

Traditionally, benthic habitats have been some of the most difficult habitats to identify and map correctly, since they involve classifying habitats under water (Diaz et al., 2004). Although they are ecologically very important to coastal ecosystems, historically, very little attempt has been made to create quantitative maps of these areas. One reason for this lack of mapping may be that benthic habitat areas tend to be quite small and relatively difficult to see with the naked eye when compared to land habitat types. However, remote sensing has proven to be one of the most effective and efficient methods for mapping land cover types over larger land based areas so as remote sensing techniques have improved in the past few decades, researchers proposed that this method could also be very useful for mapping benthic habitats (Macleod and Congalton, 1998).

In the past, a primary source of remotely sensed data has been analog aerial photographs (Congalton and Green, 1992; Ferguson and Korfmacher, 1997; Jensen, 2005; Congalton and Green, 2009). However, with the advent of digital cameras, a new and very useful form of these aerial images comes in digital format, ready to be used in a Geographic Information System (GIS). Digital data in the form of satellite images has been used for decades, but until recently the spatial resolution of satellite imagery has never been comparable to that produced by aerial photography. Even today, digital aerial imagery is especially important in benthic habitat mapping because it has more flexibility in its temporal, spectral, and spatial resolutions than satellite imagery (Ferguson and Korfmacher, 1997; Mumby et al., 1997; Jensen, 2005). The adaptable nature of digital aerial imagery data has allowed researchers to begin to classify benthic habitats more quickly than ever before, leading to advances in classification techniques of the remotely sensed data (Friedlander et al., 2007).

In computer based classifications, there are commonly two different ways for a computer to analyze an image for classification purposes: pixel-based and object-based image analysis (OBIA) approaches (Jensen, 2005). Pixel-based approaches classify a digital image on a pixel by pixel basis without accounting for the surrounding pixels. The OBIA approach analyzes individual pixels in the context of its surrounding pixels and groups pixels with similar properties into polygons. Image segmentation parameters are used to define how each pixel on the image is grouped into polygons. The size and shape of the resulting polygons are also determined by these parameters and they are usually defined by the producers of the map. Some researchers have argued that of these two types of analysis, the OBIA analysis more closely resembles how the human mind views images and is therefore usually the preferred method (Warner et al., 1998; Blaschke and Strobl, 2001).

After choosing an analysis method, there are then three general classification algorithms that are used to classify the image polygons or pixels: supervised, unsupervised, and a hybrid of the first two ways (Jensen, 2005; Congalton and Green, 2009). Recently, a new classification algorithm called the Classification and Regression

Tree (CART) technique has been widely accepted as a very effective method for classifying remotely sensed data. This method uses a decision tree created using training points to define what characteristics can be used to classify each unidentified polygon (Hansen et al., 1996; Friedl and Brodley, 1997; De'ath and Fabricius, 2000).

As with any remote sensing project, the accuracy of the maps resulting from classification is of upmost importance. Since mapping benthic habitats using remote sensing is a relatively new endeavor, testing each approach using accuracy assessment processes is necessary. Generally, the accuracy of a map is measured in two ways, positional accuracy and thematic accuracy (Congalton and Green, 2009). Positional accuracy is the accuracy of a point on the map relative to the actual physical location of the point on the ground. Thematic accuracy refers to the label (e.g. habitat, land cover, vegetation, etc.) given to a place on the map versus the actual label found for that same place on the ground. Both are important when mapping an area, although most present day research focuses primarily on thematic accuracy since technologies used to register images to the ground have vastly improved, minimizing positional error.

With advances in technology come new problems to solve. This research looks at issues associated with the classification of digital aerial imagery for the purpose of mapping benthic habitats. Remotely sensed data of the study area were initially acquired using a digital aerial camera and then classified using an object-based approach. Since an object-based approach was used, the segmentation parameters used in this study were investigated as one of the processes that may have had an effect on the overall quantitative accuracy of the final map. Other issues investigated as part of this research were the specific placement of the accuracy sites used in the accuracy assessment of the map and other common difficulties associated with classifying benthic habitats, such as the separation and identification of habitat types, and how these issues might affect the overall accuracy of the classification process and the resultant map.

History of Mapping Benthic Habitats

According to the National Oceanic and Atmospheric Administration (NOAA) (NOAA Coastal Services Center, 2008), "benthic habitats can best be defined as bottom environments with distinct physical, geochemical, and biological characteristics," or any environments found at the bottom of a water body. Marine benthic habitats are split into five distinct zones according to their depth below water level. The deepest zone is the hadal zone at over 6,000 meters deep. The abyssal zone ranges from 2,000 to 6,000 meters and the bathyl zone is between 200 to 2,000 meters deep. The shallowest zones are the nearshore and the estuarine zones, both at less than 200 meters deep (NOAA Coastal Services Center, 2008).

These two shallow zones, nearshore and estuarine, are the most researched marine zones since they are also the most accessible. The benthic habitats in nearshore and estuarine zones are also the most important when researching coastal marine environments since they account for much of the nursery, foraging grounds, and protective habitat for many fish species. In addition, they can be used as a very good indicator of the overall health of the estuarine ecosystem. The nearshore benthic habitats also dispel wave energy to provide a buffer for the shoreline and improve the local water quality through filtration (Macleod and Congalton, 1998; Handley et al., 2007).

The importance of these habitats in maintaining healthy coastal ecosystems made it increasingly clear that accurate maps for these areas are necessary. Having up to date and continuous maps of the nearshore zone can improve resource management as well as aid in monitoring the effects of humans on benthic habitats (Ferguson and Korfmacher, 1997; Mumby and Harborne, 1999, Bostrom et al., 2006). These maps can be used for anything from change detection and conservation decisions to the maintenance of shipping routes. Historically, the only way of creating maps that resemble the actual coastal benthic habitats was to conduct very lengthy and labor intensive site visits to the area and count the number and types of habitats present. This method was very imprecise and subjective due to the complete reliance on the knowledge and perspective of the scientist viewing the sea floor (Mumby and Harborne, 1999). However, as the use of aerial analog and digital imagery became more prevalent, different ways of using remote sensing to map benthic habitats were explored.

The initial means of remotely sensing underwater habitats was to take analog aerial photographs of the habitats at optimal times during the day and year. A low sun angle can cause unwanted glint or shadows in an image, so an image taken with the sun directly overhead is ideal for benthic habitat mapping. Tides can also affect how an image of benthic habitats is captured. Since the habitats are under water, the lowest tide near midday of the year is preferred so that there is the least amount of water covering and possibly obscuring the benthic habitats.

Benthic habitats can be some of the hardest habitats to image. The sun angle and tide level requirements for taking benthic habitat images may seem restrictive, but there are also several other factors that can make taking the image nearly impossible. For example, rough water caused by inclement weather can reflect light in various directions making capturing an image very difficult. Turbid water can also scatter light before it reaches the seafloor, therefore resulting in images of murky water without capturing any detail of the benthic habitats.

Even if all of the conditions are perfect for aerial photography, the imagery can still only penetrate the water up to 25 meters in depth, limiting imagery to specific areas of the nearshore and estuarine zones (Mumby and Harborne, 1999). These limitations are applicable to analog photography as well as digital aerial imagery and satellite imagery. However, since digital imaging technology is continuously improving, the most detailed imagery of benthic habitats is now usually created using digital aerial cameras rather than their analog counterparts. This change is primarily because digital imagery is now able to capture the spectral reflectance of benthic habitats in a number of different wavelengths (i.e. bands) at a fairly high resolution, giving researchers considerably more information about the reflectance of the habitats than they would receive with analog photography, which is limited to 3 bands of light.

In order to supplement imagery, Light Detection and Ranging (LIDAR), Radio Detection and Ranging (RADAR), and boat mounted Sound Navigation and Ranging (SONAR) technologies have been used to sample the floor structure and give observers more clues as to the habitats present in the nearshore and estuarine zones (Mumby and Harborne, 1999; Zajac et al., 2003). These sensing techniques do not produce images as analog or digital cameras would, rather they are active sensors. Active sensors, such as LIDAR or SONAR, produce the energy that they then reflect off the structure they are trying to sense (e.g. the seafloor). Differences in return times of the emitted energy are then used to calculate distances to that structure. From these distances researchers can make inferences about the structure of the seafloor. These technologies have also been used to remotely sense habitats at greater depths, but with less precise classifications since only the structure type is sampled, rather than the spectral reflectance of each habitat type.

Classifying Benthic Habitats

There are generally three label types that researchers choose from when designing a classification scheme to classify benthic habitats. The first label type labels the habitat using its geomorphology, which is the general landscape of the area (e.g. spur and groove zone). The second label type labels the physiognomy of the habitat, which can be defined as the outer appearance of the habitat (e.g. coral reef). The third label type labels the habitat according to its ecology, which is the label given to the organism and the environment (e.g. turf algae). In addition, sometimes a geological history label is given to a habitat, which defines the geologic structure of the area (e.g. relic reef) (Mumby and Harborne, 1999). To ensure that the classification scheme labels each habitat precisely and accurately, one of these label types should be chosen to be used for the entire benthic habitat mapping project. On occasion, a combination of these labels may be used, but this amalgamation of labels often leads to more confusion than information about the area.

In the past, the majority of benthic habitat classification was done by site visit (Macleod and Congalton, 1998). This time-consuming endeavor created very coarse and discontinuous maps that were rarely updated or redone. The process of mapping by site visit was not conducive to change detection mapping and, in actuality, very little total area was classified using this method. Then, between the late 1980s and 2002, remote sensing of benthic habitats became more commonplace, but classification schemes were not yet standardized (Madley et al., 2002). In many cases, habitat label types were mismatched or misused, making it nearly impossible to compare two or more maps. During this time, databases for general benthic habitat maps were in their beginning stages and only very limited areas had meaningful coverage (Mumby and Harborne, 1999).

Since 2002, researchers have become more proficient at creating usable benthic habitat maps that have some consistency in classification. This consistency will allow these new maps to be compared with other maps that were created using the same general method. Once the initial imagery has been properly collected and prepared, current classification methods for creating benthic habitat maps generally involve these main steps: 1) creating a comprehensive classification scheme; 2) choosing and implementing a classification algorithm; and 3) performing an accuracy assessment on the classified map (Congalton and Green, 2009).

Classification Schemes

A classification scheme, or system, is a way of consistently placing specific map data into more useful labels or categories. For instance, a classification scheme can take the relatively hard to comprehend raw number data of digital imagery and categorize areas into functional labels, such as "water" or "land". According to Congalton and Green (2009) each classification system should have well defined labels that are totally exhaustive, mutually exclusive, and hierarchical. The first two criteria guarantee that every section of the given image fits into a single category with no overlapping categories and no areas left unclassified. The third criterion ensures that the classification scheme is made up of levels of classification, from very broad categories to much more specific labels, so the classification can be as detailed as necessary for the study at hand. Since each classification scheme is usually chosen or created based on the land cover types present in the study area, it is very important that each scheme have exact definitions for each of its chosen labels and a set of rules for assigning each area to a label. Without specific definitions and rules for each of the labels, the application of a classification scheme would be random and inconsistent throughout the image (Congalton and Green, 2009).

Before 2002 there were several schemes in use for classifying benthic habitats. The most prevalent were the systems created by the US Geological Survey (USGS) (Anderson et al., 1976), US Fish and Wildlife (Cowardin et al., 1979), and the NOAA Coastal Change Analysis Program (Dobson et al., 1995). In 2002, a group of scientists researching benthic habitats on the Florida coast proposed taking the many available classification schemes and combining them into one broad, but usable scheme (Madley et al., 2002). The three most widely recognized schemes noted above and two Florida specific schemes from the Florida Department of Transportation (1985) and the Florida Natural Areas Inventory (1990), were all drawn upon to create the Florida System for Classification of Habitats in Estuarine and Marine Environments, or SCHEME (Madley et al., 2002). In general, SCHEME is very similar to any good classification system in that it is well defined, totally exhaustive, mutually exclusive, and hierarchical (Appendix

A). However, unlike many systems, SCHEME is very dynamic in nature and is defined by its ability to have categories added or subtracted from the overall scheme as necessary. This capability makes it possible for SCHEME to be used in conjunction (i.e. crosswalk) with various other classification systems.

Although SCHEME was created for use in Florida, it has become a very popular classification system for benthic habitat mapping all over the globe, in a large part because it creates a useful standard for classification. Maps that are created using SCHEME can be compared over both time and location, because the habitat labels are standard but can be modified or changed according to specific site needs. Since benthic habitats in different conservation areas may be linked through tides, having a classification system that can be used in any region is a very important advantage in mapping (Mumby and Harborne, 1999). Also, the quantitative descriptors used in SCHEME give more consistency to labeling habitats, thus making the understanding and comparison of maps created by different producers much easier. Overall, SCHEME is a system that allows anyone to create a usable benthic habitat map that can be compared with almost any other benthic habitat map produced using the same or a similar scheme, thus making database creation and maintenance for each study area much simpler.

Classification Algorithms

The objective of a classification algorithm is to link changes in spectral reflectance depicted on the imagery with actual physical changes in habitat found on the ground (Congalton and Green, 2009). With a good classification scheme and a well executed classification process, habitats that are the same on the ground and appear similar on the image will be placed into the correct category and given the same label. As stated above, there are three general categories of classification algorithms: supervised, unsupervised, and hybrids of both (Jensen, 2005). Each of these three types of algorithms can be applied to the remotely sensed data in various ways.

Supervised Algorithms

In a supervised classification algorithm, training sites on the imagery are given a label prior to classification based on knowledge collected through a combination of fieldwork, photo-interpretation, personal familiarity with the site, and various other sources (Jensen, 2005). These training sites are usually collected in areas of homogeneous habitat so that the spectral properties of these areas are unique to a specific habitat. The typical recommendation for images with medium to low spatial resolution is to use homogeneous areas of at least 3 pixel sizes in length and width to ensure that the training point is located in the correct habitat on the map even with the minimal positional error introduced by the global positioning systems (GPS) used to collect the data (Jensen, 2005; Congalton and Green, 2009). However, if imagery with higher spatial resolution is used, a 3 by 3 pixel box of homogeneous area may not be sufficient for taking training sites. For instance, if an image with 2 meter pixels is registered to the ground with a positional error of 8 meters, a 6 by 6 meter homogenous area (i.e. 3 pixels in height and width) would not be big enough ensure that the training point would fall within this habitat once located on the map. In this case a larger homogenous area, possibly a 5 by 5 pixel box, would have to be used (Congalton and Green, 2009).

After collecting training points, the remaining unknown sections of the image (those that were not identified as training points) are statistically compared with the chosen training sites. The unknown sections are then grouped into the same classification as the training point to which they are most similar. There are generally two types of statistical comparison techniques for classifying unknown sections of images using supervised classification algorithms: nonparametric and parametric. For the purposes of this discussion, the classification format will be assumed to be a pixel-based approach; however, the same statistical techniques can be used for object-based approaches. Parametric classification techniques assume that the data are normally distributed, while nonparametric techniques do not (Jensen, 2005). The most prevalent nonparametric techniques include: parallelepiped, minimum distance, and nearest neighbor. The maximum likelihood algorithm is the most widely used parametric technique.

The parallelepiped method uses the training sites to define decision boundaries for each band (or specific wavelength of spectral reflectance) captured by the imagery (Figure 1) (Jensen, 2005). The decision boundaries are the upper and lower brightness value limits in each category as defined by the highest and lowest values found in the training sites of each category. If an unknown pixel falls within a set of defined decision boundaries, it is given the label of the category with those limits. An advantage to using this method is that the distribution of the training sites in each category is taken into account without the need to do any lengthy calculations. However, the drawback to this method is that in some instances a pixel may fall within two categories or none at all, in which case it is given an unclassified label.

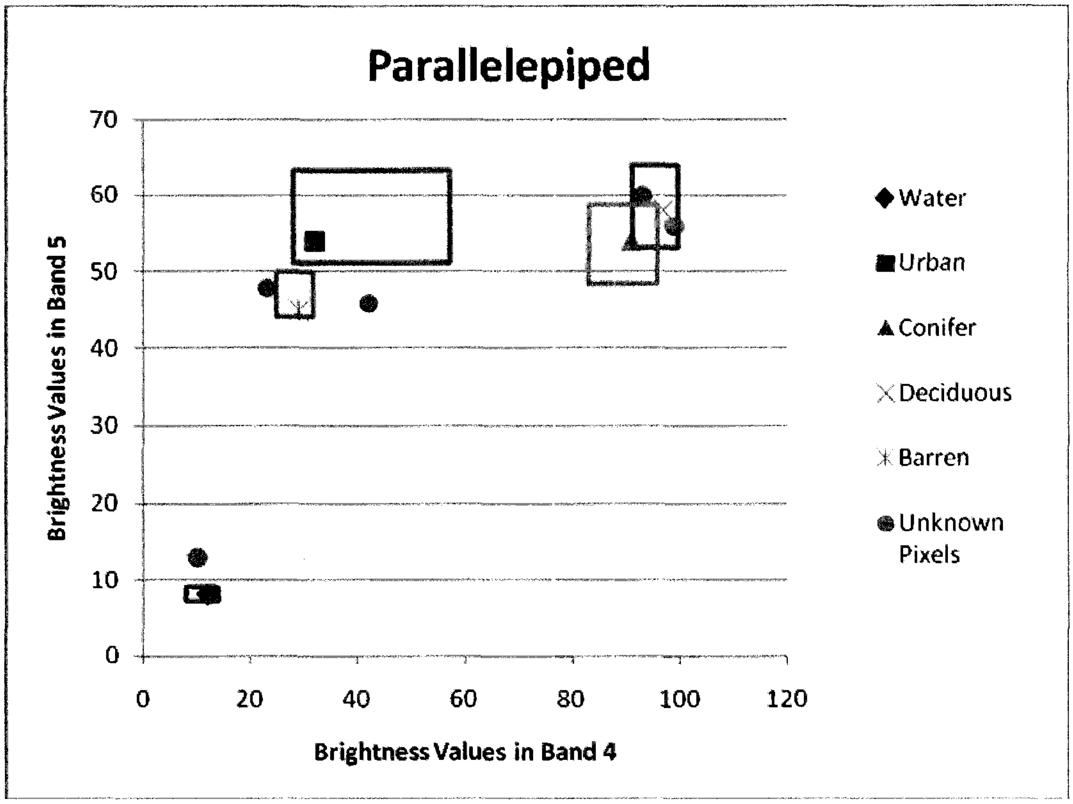


Figure 1: An example of the parallelepiped technique used for land cover types. The markers represent the mean values of the training sites used to define each category and the matching boxes represent the decision boundaries for those categories. The orange circles are unknown pixels. Using this classification the two orange circles inside the deciduous forest boundaries would be classified as deciduous and the other three would be labeled unclassified.

The minimum distance and the nearest neighbor techniques are very similar to one another, and both rely on calculating similarities in spectral reflectance between an unknown pixel and a known training site. Calculating the similarity in spectral reflectance doesn't require any more than a simple Euclidean distance calculation between points when spectral reflectance is plotted on a brightness graph as depicted in Figure 2. The minimum distance technique classifies unknown pixels based on the category with the closest mean brightness values (calculated across all training sites in that category), while the nearest neighbor technique classifies unknown pixels based on the training site with the closest brightness values (Jensen, 2005). Neither of these techniques account for the distribution of the training sites within categories which can lead to significant misclassification. For instance, it may be possible that when using the minimum distance technique an unknown pixel is closest to the mean of a very closely populated "Barren" category while with the nearest neighbor technique the nearest training site is in the more spread out "Urban" category (Figure 2). In most instances, the minimum distance technique is quicker and easier than the nearest neighbor technique since it only calculates the distance to the means of categories rather than to every single training point. However, both of these techniques ensure that every unknown pixel is given a specific category other than unclassified, which can be very advantageous during classification.

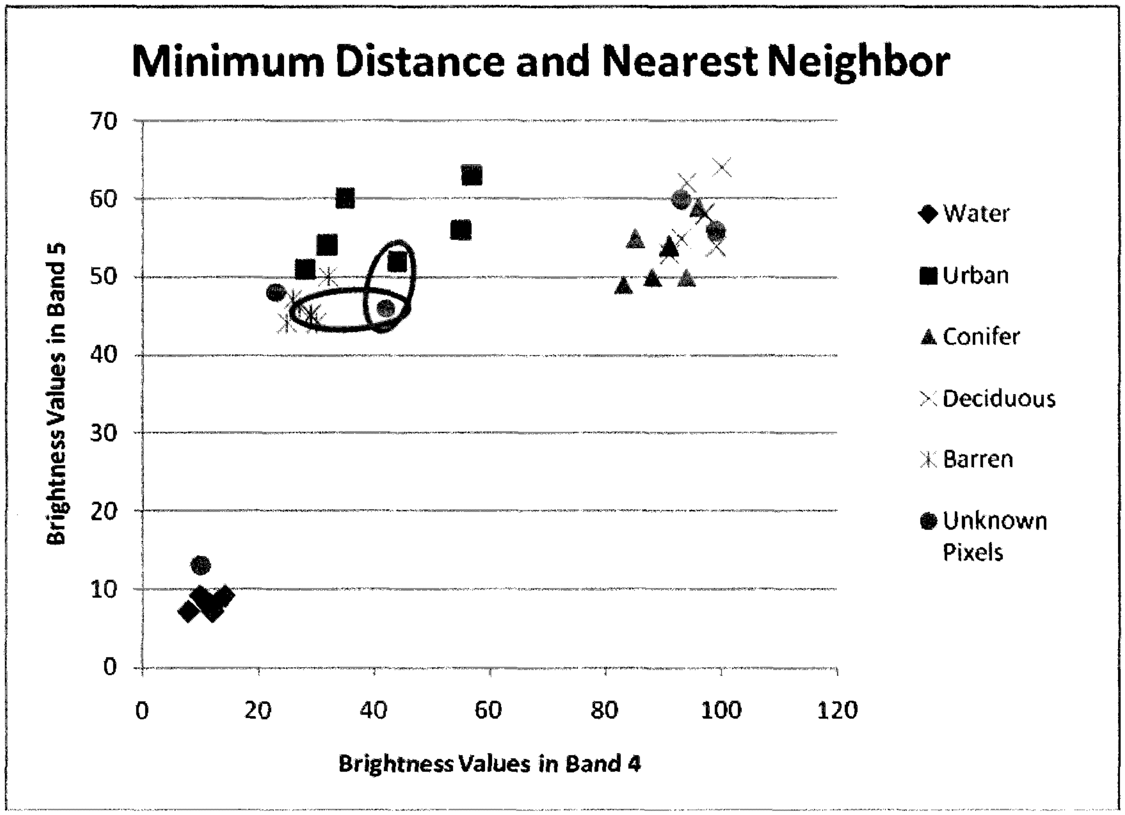


Figure 2: An example of the minimum distance and nearest neighbor techniques. Each category has five training points and a mean value shown in black. The orange circles represent the unknown pixels. The blue circle shows that in the minimum distance technique that specific unknown pixel would be classified as barren while in the nearest neighbor technique the same pixel would be classified as urban as shown with the red circle.

Of the four most prevalent statistical techniques mentioned above, the maximum likelihood analysis requires the most complex calculations since it is probability based rather than distance based. The maximum likelihood classification technique computes for each unknown pixel the probability of it belonging to each category and then chooses the category with the highest probability (Jensen, 2005). The probabilities are based on the mean brightness values and the variances of the categories. Therefore, this technique accounts for the distribution of training sites in each category. Another advantage to this technique is that it classifies every unknown pixel, since each one receives a probability for each and every category, no matter how small the probability may be. However, this technique assumes that the data are normally distributed, which may not always be true. Therefore, this technique is not always applicable (Jensen, 2005).

Unsupervised Algorithms

In an unsupervised classification algorithm the unknown pixels are grouped into clusters with similar spectral properties without using any ancillary data (Jensen, 2005). The lack of training data means that the beginning of this classification process has very little human involvement. The analyst typically has to define the number of final clusters they want the computer to place the unknown pixels into, the maximum number of iterations that will be used to swap the pixels to minimize the variance within the clusters, and the convergence threshold or the maximum percentage of unchanged pixels between iterations (Chuvieco and Congalton, 1988). Once all of the unknown pixels are placed into clusters, photo-interpretation or training sites can be used to combine and label clusters into meaningful classification categories, such as "water" or "land". An advantage of this classification technique is that the computer may be able to pick out subtle changes in land cover types that a supervised classification might overlook. Another advantage is that training data doesn't need to be collected until after the classification is complete and therefore can be selectively collected in areas that help to define certain clusters.

The most prevalent unsupervised classification algorithm is the Iterative Self-Organizing Data Analysis Technique (ISODATA) (Jensen, 2005). ISODATA is a clustering program that starts by placing the pixels arbitrarily into the chosen number of clusters and then comparing each pixel's brightness values with the computed cluster mean values. The program then shifts the pixels into more appropriate clusters so that the variance between clusters is greater than the variance within the clusters. It then recalculates all of the cluster mean values. ISODATA repeats the shifting of pixels until either the maximum iteration number or the convergence threshold has been met. Since the initial placement of the pixels is random, the creation of the clusters is free of human biases (Jensen, 2005). After the process is complete, the analyst can then use training data to aid in the labeling of the clusters.

CART and Hybrid Algorithms

While both the supervised and unsupervised classification algorithms can be quite successful, many have found that a hybrid of the two classification types yields even better results (Chuvieco and Congalton, 1988; Hansen et al., 1996; Macleod and Congalton, 1998). Many hybrid classification algorithms exist, such as cluster-busting or artificial neural networks. However, another prominent classification algorithm is the Classification and Regression Tree (CART) technique. This technique is nonparametric, so it therefore makes no assumptions about the distribution of the raw data (Hansen et al., 1996; Friedl and Brodley, 1997).

In most instances the CART technique is considered a supervised algorithm, but in many ways it shares characteristics with hybrid algorithms. The CART technique uses initial training data with labels to "grow" a hierarchical decision tree, which can then be used to classify unknown pixels. The decision tree is created using a binary partitioning algorithm that defines independent predictor variables to split the initially large heterogeneously labeled group of training data into smaller groups with unique spectral and informational properties. At each branch of the decision tree, the algorithm picks the variable that best divides the data into separate categories, and it continues to split the data using new predictor variables until it has split the data into final single category groups (Hansen et al., 1996). The final homogeneous groups are often referred to as end nodes and there is usually more than one end node per category. This splitting of the categories by both spectral and categorical properties gives the CART technique somewhat of a hybrid feel, especially when it is used in conjunction with an OBIA approach, which groups pixels solely based on their spectral properties.

The decision tree created by the binary partitioning algorithm can then classify any unknown pixel into a category by sending the pixel through the tree. The unknown pixel follows the decision tree through the many different predictor variables and if the pixel is consistent with the variable, the pixel proceeds down the right branch of the tree. However if the pixel is inconsistent with the variable, it proceeds down the left branch of the tree. Once the classification is complete, the predictor variables used to classify each pixel can then tell the user a lot about the nature of the pixels (Hansen et al., 1996).

For example as viewed in Figure 3, an initial predictor variable might be whether the brightness value of the pixel is greater than 100 for the blue band of the image. If the pixel does not have a brightness value greater than 100 for its blue band, then it follows the left branch of the tree to the next predictor value. Then, if the next predictor variable determined whether the pixel had a value higher than 100 in its near infrared band (NIR) and it did not, it would continue through the left branch again to the predictor variable that determined whether it had a green value less than 100. If the answer to this predictor variable was again no, the pixel would be classified as "Grass" (Figure 3).

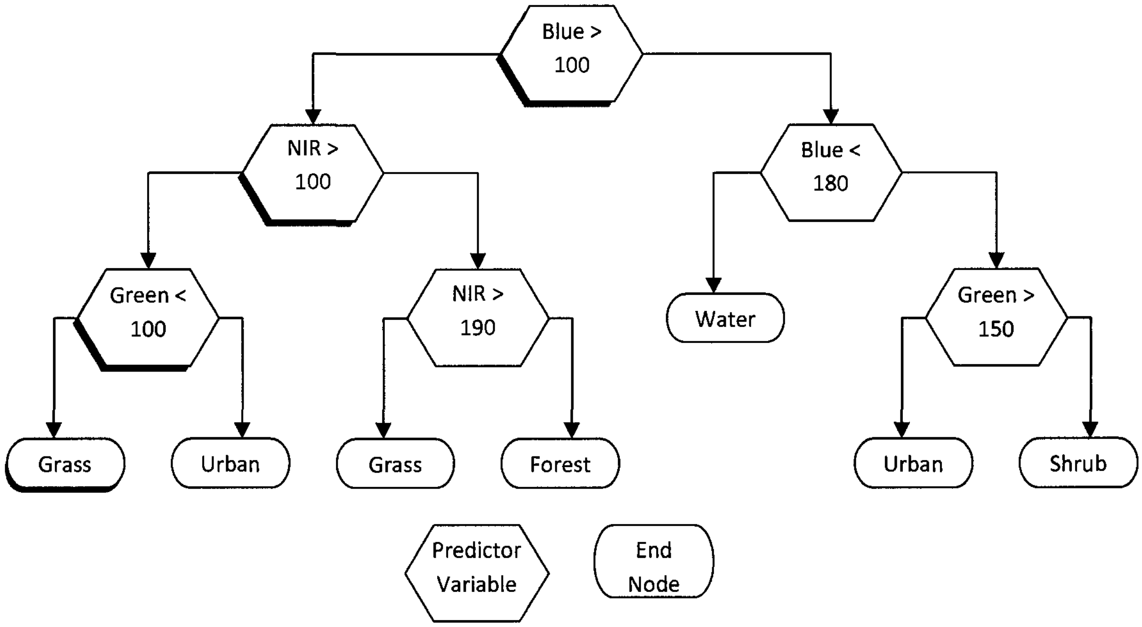


Figure 3: An example of a decision tree used in the CART technique. Note that there are multiple end nodes for the same classification categories and predictor variables are independent of any other variables.

Object-based Image Analysis

As stated previously, all of these classification algorithms can also be applied to images using an object-based image analysis (OBIA) approach. In the OBIA approach unknown pixels are grouped into polygons and then classified as polygons rather than individual pixels, as seen in the pixel-based approach. Classifying pixels as a group, rather than individually, actually increases the number of attributes that can be used to identify the pixels, such as polygon shape and texture (Baatz et al., 2001). Unlike a single square pixel, a polygon has a given shape that can give some insight into the habitat type contained within that polygon (e.g. a rectangular building vs. a round swimming pool). Since many pixels within a polygon will not have exactly the same spectral properties, the texture created by the variations within the polygon can also be indicative of the enclosed habitat (e.g. a mixed forest stand vs. a grassy field). This extra information, added by choosing to use an OBIA approach to classify an image, can vastly increase the accuracy and effectiveness of the mapping technique.

Although there are many advantages to using the OBIA approach over the pixel-based approach, there are some caveats when dealing with polygons rather than individual pixels. In OBIA classification approaches the image is broken into polygons of similar unknown pixels so that there is less spectral variation within each created polygon than between the polygons. This grouping of pixels is called segmentation. However, in the majority of images, the pixels that comprise one habitat type do not all share exactly the same spectral properties, meaning the segmentation may or may not place these pixels into the same polygon (Blaschke and Strobl, 2001). Therefore, some decisions need to be made to determine how the image will be segmented.

One of the most common and robust programs used to segment images is Definiens' eCognition (Baatz et al., 2001). This program, like many others, uses a bottom-up approach to segmentation, meaning a set of parameters is defined by the researcher so that the program can then automatically segment the entire image using these set parameters. The researcher defined parameters are: layer weights; image object level; segmentation mode; scale parameter; composition of the homogeneity criterion; and type of neighborhood (Baatz et al., 2001).

The first segmentation parameter determines how much weight is given to each spatial layer of information being used in the classification. The weights of all of the layers should sum to 1. The second parameter can be defined if the researcher wishes to have more than one level of segmentation. For example, the researcher may want to complete a very general segmentation resulting in very large polygons and then a more specific segmentation, nested within the first segmentation, to break up the larger polygons into more specific smaller polygons. Each of these segmentations would then be given a specific level. The segmentation mode parameter is usually set to normal (i.e. just defining polygons), but can be set to create sub-objects, or polygons found within the context of other polygons. The mode parameter can be used in conjunction with the object level parameter so that sub-objects are defined as a new level.

The scale parameter is one of the most important segmentation parameters defined by the researcher. This parameter determines the maximum allowed heterogeneity within each of the created polygons. A higher number results in larger polygons with a larger variety of pixels, while a lower number results in smaller polygons with a more homogeneous makeup of pixels. Therefore, this parameter affects how well the segmentation breaks up like and dislike pixels, and can vastly change the resulting map.

Traditionally, the thought was that the smallest segmentations classify an area with the most detail; however, more recent research has found that the smallest segments aren't always the most accurate (Schiewe, 2002; Rahman et al., 2003). In many instances a low scale parameter will create two polygons out of the same object. For instance, if an image is taken of a tree while the sun is not directly over the tree, half of the tree's canopy will be in shadow, giving it slightly different spectral properties than the lighted half. If the image is then segmented using a very small scale parameter, half of the tree's pixels may be placed in one polygon and the other half in another. This segmentation may lead to half the tree to being classified in one category, and the other half of the tree being classified in another. The smaller segmentation parameter therefore causes more error in the classification process than if the initial scale parameter had been slightly larger, allowing the tree to be segmented into one polygon instead of two.

The composition of the homogeneity criterion deals with the parameters regarding how much emphasis should be placed on the color and the shape of the object. Generally, color is very important when segmenting an image and is usually given the most attention (Baatz et al., 2001). However, shape can also help define certain objects. The shape criterion is actually divided into two different categories: smoothness and compactness; and each of these criterions can be weighted along with color to determine how much each category should be considered during segmentation.

Lastly, the type of neighborhood parameter is used to determine whether the diagonal pixel neighborhood technique should be used while segmenting the image. This technique allows pixels that share a corner (rather than a full side) to be considered one continuous polygon. When the diagonal pixel neighborhood technique is not used, only the 4 adjacent pixels (that share a side with the pixel in question) are considered in that pixel's neighborhood during segmentation. Therefore, only the 4 adjacent pixels can be segmented into the same polygon as the original pixel and a new polygon is made out of the diagonal pixels, even if they have the same label (Figure 4). However, if the diagonal pixel neighborhood technique is used, all 8 of the adjacent pixels to the original pixel are considered within the pixel's neighborhood and can be segmented into the same polygon (Figure 4). This ability is really only powerful when objects within the image are about the same size as the pixels. For instance, this technique would allow a road with a width that is less than or equal to the size of one pixel in the image to be defined as one continuous road, rather than a few diagonally distributed pixels of impervious surface.

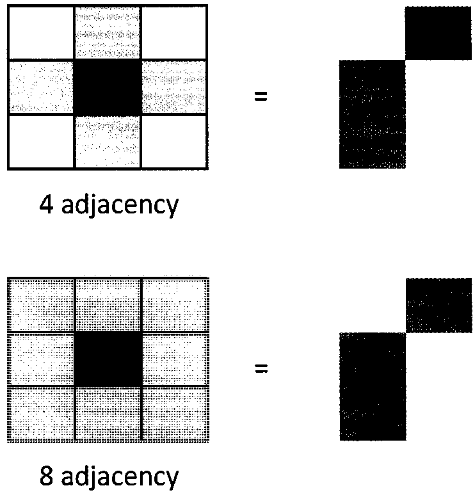


Figure 4: The difference between 4 adjacency and 8 adjacency. The original pixel is in black and the neighborhood of the original pixel is shown in yellow. Any pixels within the neighborhood of the original pixel can be segmented into the same polygon as that pixel. In 4 adjacency, pixels that share a side become one polygon (shown in blue) and diagonal pixels become another polygon (shown in red). In 8 adjacency (i.e. when using the diagonal pixel neighborhood technique), diagonal pixels may form a polygon with the original pixel.

Once the image is successfully segmented, the polygons can then be classified using any of the techniques discussed above for the pixel-based approach. However, when collecting the training sites to use during classification, the size of the segmentations can become important. As with the pixel-based approach, training points are usually taken in areas of homogeneous habitat that are large enough so that the training points always fall within that homogeneous habitat even with the computed positional error. Often the minimum mapping unit (mmu), or the smallest acceptable polygon size for the image, is large enough to account for any positional error associated with collecting and registering the training point. In these cases, the training points can be collected within an area the size of the mmu and then the information can be tied to the polygon in which that point resides. When the mmu is still smaller than the expected positional error, a group of polygons can be used as the homogeneous habitat area needed for a training site, but this is the less preferred method. In most cases the mmu should be large enough to account for the positional error.

Accuracy Assessment

In order to evaluate how well a classification is carried out, an accuracy assessment is performed on the resulting map. The accuracy assessment helps to identify where there might be errors in the map and provides users with the overall reliability of the map (Congalton, 1991; Jensen, 2005; Congalton and Green, 2009). Accuracy assessments generally involve a number of individual accuracy sites which are used to represent the entire map. The classification labels of these accuracy sites are compared to the reference data labels of the same site. The reference data refers to the data assumed to be correct, or at least the standard for comparison for the map. The comparison of the classification labels to the reference data is usually summarized in an error matrix to quantitatively depict where errors occur in the map. After the error matrix is generated, various other statistical tests can be performed to further investigate the accuracies and the effectiveness of the maps. Two of the more prevalent statistical tests are the Kappa and MARGFIT analyses (Congalton, 1991; Jensen, 2005; Congalton and Green, 2009).

History of Accuracy Assessment

Although remote sensing and map making are relatively old endeavors, accuracy assessment is still a relatively new development, and only really started to become a priority over the past few decades. Beginning after World War II, photo-interpretation was used to create maps from aerial photography. However, at that time there was no way of quantifying the accuracy of these maps. Most of the time, if the map "looked good", or looked like it represented the aerial photo, then it was assumed to be accurate (Congalton and Green, 2009). However, between the 1950s and the 1970s many researchers expressed the need for a better way to assess the reliability of maps, but it wasn't until after the launch of Landsat 1 in 1972 that any real progress was made to accomplish this task (Jensen, 2005; Congalton and Green, 2009).

Early versions of an accuracy assessment started to appear in the late 1970s and were usually non-site-specific assessments. The term "non-site-specific assessment" refers to the process of comparing the overall areas of the land cover types on the map with the assumed total areas of each land cover type on the reference data. However, at that time there was no definition for what constituted acceptable reference data. When attempting to assess the accuracy of a map generated from early satellite imagery, the map was often compared to a reference map that had been photo-interpreted from aerial photos. The reference map often had only been assessed by the "looks good" technique, meaning the map appeared to look similar to the aerial photo. In these instances, the accuracies of the satellite-derived maps really had no meaning and were only relative to the maps being used as reference data (Congalton and Green, 2009). Even if the reference data were perfect, this assessment still did not address the location of the land cover types. This disregard for location meant that the same amount of a specific land cover type could be present in the newer classified map, but in a different place than the reference map, and the accuracy assessment of the classified map would not address this inconsistency (Congalton, 1991; Congalton and Green, 2009).

In order to get a clearer picture of where classification algorithms may cause error, site-specific accuracy assessments were introduced in the 1980s. This new process not only assesses the accuracy of the total number of labels given, but also the accuracy associated with position of the given labels (Congalton, 1991; Congalton and Green, 2009). The assessment addresses whether sites are not only correctly labeled but also in the correct place. Because the site-specific assessment produces a more detailed accuracy estimate, it quickly became the preferred accuracy assessment technique and an error matrix was fashioned as a standardized way to present the findings of accuracy assessments (Congalton, 1991; Jensen, 2005). An error matrix, also referred to as a confusion matrix or contingency table, not only calculates the overall site-specific accuracy of each category used in the classification, but also how often one category is mistaken for another category (Jensen, 2005; Congalton and Green, 2009). The error matrix also serves as a jumping off point for many statistical techniques that may be used to better summarize the accuracy assessment (Congalton, 1991).

With the growing use of the error matrix, concern over the acquisition of reference data grew. Congalton (1991, pg. 42) expressed that "Although no reference data set may be completely accurate, it is important that the reference data have high accuracy or else it is not a fair assessment. Therefore, it is critical that the ground or reference data collection be carefully considered in any accuracy assessment." This statement indicates that not only the accuracy of the final map is of importance, but also the accuracy of the reference data. The "looks good" method is no longer good enough. In many cases, this means that the reference data are collected by physically visiting the accuracy sites in the field to determine their validity. However, visiting each accuracy site can be very costly and time-consuming, and not necessarily 100% correct because of observer error or unaccounted for changes due to time between the acquisition of the remotely sensed data and the reference data. Therefore, photo-interpretation can still be used to collect some of the reference data. The difference in the current photo-interpretation collection methods, as compared to previous methods, is that the accuracy of the reference data is now closely monitored, whereas before it was almost completely ignored (Congalton, 1991).

Error Matrix

The error matrix is a square array of numbers in rows and columns that represent the number of accuracy sites placed in each category for the classified map data as compared to the reference data (Table 1) (Congalton et al., 1983; Congalton, 1991; Jensen, 2005). The columns represent the reference data being used to check the accuracy of the map and the rows represent the map data. For example, if an accuracy site on the map is identified as "Barren", but in the reference data it is labeled as "Agriculture", that site would be one of the three tallies placed in the second box down in the farthest left column of Table 1. The major diagonal represents total agreement between the map data and the reference data. The overall accuracy is calculated by summing the major diagonal and dividing it by the total number of accuracy sites, as shown at the bottom of Table 1 (Story and Congalton, 1986). Ideally, a perfect map would have all zeros in the non-major diagonal boxes, meaning no accuracy sites were misclassified (Congalton et al., 1983).

Table 1: A sample error matrix.

Map Data	**Reference Data** Agriculture	Barren	Forest	Lake	**Row Total**
Agriculture	11	0	0	0	11
Barren	3	17	0	0	20
Forest	6	3	18	4	31
Lake	0	0	2	16	18
Column Total	20	20	20	20	80

Producer's Accuracy			**User's Accuracy**		
Agriculture	11/20	55%	Agriculture	11/11	100%
Barren	17/20	85%	Barren	17/20	85%
Forest	18/20	90%	Forest	18/31	58%
Lake	16/20	80%	Lake	16/18	89%

Overall Accuracy	11+17+18+16= 62	62/80	**78%**

The error matrix is an effective way of depicting not only the overall accuracy, but also to evaluate errors of inclusion and exclusion (Congalton et al., 1983; Congalton, 1991; Jensen, 2005). Exclusion or omission errors are calculated by dividing the number of correctly classified accuracy sites by the total number of accuracy sites collected for that category (i.e. the column total for that category). The percentage produced using this process represents the probability that an accuracy site of this reference label will be correctly classified in the map (Story and Congalton, 1986; Jensen, 2005). These percentages are most commonly referred to as "producer's accuracies" since from them a producer of a map can tell how well a certain land cover category can be classified using their classification method (Story and Congalton, 1986; Congalton, 1991). On the other hand, errors of inclusion, or commission errors, are calculated by dividing the number of correctly classified sites in a category by the total number of sites classified in that category (i.e. the row total for that category). The percentage produced represents the probability that the accuracy site on the map was classified correctly, matching the reference data (Story and Congalton, 1986; Jensen, 2005). These percentages are most commonly referred to as "user's accuracies" since a user of the map can use these numbers to estimate how reliable the map is for each category (Story and Congalton, 1986; Congalton, 1991).

When choosing accuracy sites to be used in an error matrix, many of the same principles are used as when collecting training sites. In many cases, both types of sites can be collected at the same time and later split into training and accuracy sites. Most projects usually reserve a minimum of 30 accuracy sites per land cover category. However, for a more effective accuracy assessment, anywhere from 50 to 100 accuracy sites for each category are recommended (Jensen, 2005; Congalton and Green, 2009). As with collecting training sites for a pixel-based classification, the accuracy site should be located in a homogenous land cover type at least as big as a 3 x 3 pixel square or larger depending on the spatial resolution and the positional accuracy associated with the data. The same is true for the OBIA approach, except the mmu and the size of the polygons should also be considered when choosing accuracy sites. Locating an accuracy site in a group of homogeneous pixels, instead of a single pixel, helps to eliminate error caused by registration issues. If, for instance, an accuracy site is chosen as a single pixel, but the image is off by a few meters when it is rectified to the ground, the reference data label of that accuracy site might actually be tied to another pixel with a different habitat type, giving that pixel inaccurate reference data. Positional accuracy can also be an issue when collecting accuracy sites for object-based approaches, so using the minimum mapping unit for the classification may be an appropriate way to define how large an area should be in order to be considered for use as an accuracy site (Congalton and Green, 2009).

The error matrix typically assumes that an accuracy site was either correctly or incorrectly labeled during the classification process, without any middle ground. However, in nature there may not always be a clear right answer. For instance, a section of forest might be divided evenly between deciduous and coniferous trees and the classification technique might label this area as either category and be equally correct. In these cases, a more lenient system might be necessary that permits both categories to be considered accurate. An accuracy assessment that allows for more than one acceptable category to be applied to a specific accuracy site is called a fuzzy accuracy assessment (Congalton and Green, 2009). Fuzzy accuracy assessments can be represented in an error matrix with two numbers in each of the non-major diagonal boxes. The first number represents the accuracy sites that are not classified in the expected category but are still in categories that are considered acceptable. The second number represents the accuracy sites that are classified in the absolutely wrong category (Table 2). The fuzzy accuracy can be calculated as done in the non-fuzzy (i.e. deterministic) error matrix by adding the major diagonal number and dividing by the total number of accuracy sites, only the fuzzy accuracy calculation also includes the number of acceptable answers in the major diagonal totals, as demonstrated at the bottom of Table 2.

Table 2: A sample error matrix with fuzzy information.

	Reference Data				
Map Data	Agriculture	Barren	Forest	Lake	**Row Total**
Agriculture	11	0/0	0/0	0/0	11
Barren	2/1	17	0/0	0/0	20
Forest	3/3	1/2	18	0/4	31
Lake	0/0	0/0	2/0	16	18
Column Total	20	20	20	20	80

Deterministic Producer's Accuracy			**Deterministic User's Accuracy**		
Agriculture	11/20	55%	Agriculture	11/11	100%
Barren	17/20	85%	Barren	17/20	85%
Forest	18/20	90%	Forest	18/31	58%
Lake	16/20	80%	Lake	16/18	89%

Overall Determinisitic Accuracy	62/80	**78%**

Fuzzy Producer's Accuracy			**Fuzzy User's Accuracy**		
Agriculture	16/20	80%	Agriculture	11/11	100%
Barren	18/20	90%	Barren	19/20	95%
Forest	20/20	100%	Forest	22/31	71%
Lake	16/20	80%	Lake	18/18	100%

Overall Fuzzy Accuracy	70/80	**88%**

Discrete Multivariate Analysis

In some instances, calculating the user's, producer's, and overall accuracies is a comprehensive enough analysis of the classification method being used. However, in most cases the ability to determine the effectiveness of one classification technique and compare it with the accuracy results of another classification technique is important in deciding the most successful method for classifying a specific area. Discrete multivariate analysis techniques can be used to easily manipulate and compare accuracies of different classification methods (Congalton et al., 1983; Congalton, 1991; Jensen, 2005). When implementing discrete multivariate analyses for these purposes, the deterministic error matrix becomes invaluable because it is a functional representation of the accuracy assessment using discrete data. The deterministic error matrix data are discrete because either the accuracy sites were classified correctly or they weren't; there is only one right answer for each site (Congalton et al., 1983). Therefore, discrete multivariate techniques do not work with the fuzzy assessment data, since there may be more than one acceptable answer for each site.

One of the most widely used discrete multivariate analyses in accuracy assessment is a technique called Kappa (Congalton et al., 1983). Kappa is a measure of the overall agreement of the classification data with regard to the reference data, and is derived from the actual accuracy of the data minus the chance agreement of the same sites (Congalton et al., 1983, Congalton, 1991; Jensen, 2005). Although this technique may seem superfluous, since calculating the overall accuracy from the error matrix is relatively simple, Kappa values can be used to compare the classification method being used to other forms of classification. The Kappa technique calculates a KHAT value, which is an estimate of Kappa, also known as the coefficient of agreement (Congalton et al., 1983; Jensen, 2005). KHAT is computed:

$$KHAT = \frac{N\sum_{i=1}^{k} x_{ii} - \sum_{i=1}^{k}(x_{i+} * x_{+i})}{N^2 - \sum_{i=1}^{k}(x_{i+} * x_{+i})}$$

where k is the number of rows in the error matrix, x_{ii} is the number of sites in row i and column i, x_{i+} and x_{+i} are the row i and column i site totals, respectively, and N is the total number of accuracy sites used in the error matrix. The first term in the numerator is essentially the number of sites in the major diagonal multiplied by the total number of sites. The second term in the numerator is the row total multiplied by the column total for each land cover category. Since KHAT values are a measure of the actual accuracy minus the accuracy of chance (i.e. random) agreement, the values indirectly account for the error, or the off diagonal elements in the error matrix. Consequently, the KHAT coefficient of agreement will usually be slightly different than the overall accuracy (Jensen, 2005). KHAT values greater than 0.80 (i.e. 80%) represent strong agreement between the classification data and the reference data, while values less than 0.40 (i.e. 40%) represent very poor, or closer to chance, agreement (Jensen, 2005). Similarly, KHAT values can also be computed for each category individually, called the conditional KHAT value, by using the same equation as stated above without summing over every value for *i*.

KHAT values can also be used to determine how well an error matrix of one classification compares to the error matrices of other classifications. To do this comparison, confidence intervals of the KHAT statistic must be computed. Since the KHAT statistic is asymptotically normally distributed, the confidence intervals can be calculated by using the approximate large sample variance (Congalton and Green, 2009). Once the confidence intervals for the KHAT statistic are found, the error matrix for the classification can be compared with an assortment of other error matrices from other classifications with the same categories. These comparisons are done using a Z test. The easiest Z test to use is to test whether a classification is significantly better than a random labeling of the accuracy sites. This simple Z test is computed using:

$$Z = \frac{KHAT}{\sqrt{\hat{\sigma}^2}}$$

where $\hat{\sigma}^2$ is the variance of the KHAT statistic. In order to test whether two distinct classifications are significantly different, the Z test is then performed using:

$$Z = \frac{|KHAT_1 - KHAT_2|}{\sqrt{\hat{\sigma}_1^{\,2} + \hat{\sigma}_2^{\,2}}}$$

where $KHAT_1$ is the KHAT value for the first error matrix and $KHAT_2$ is the value for the second, and $\hat{\sigma}_1^{\,2}$ and $\hat{\sigma}_2^{\,2}$ are the respective variances. A Z value greater than 1.96 usually indicates a significant difference between error matrices at the 95% confidence level.

Another way of comparing classifications is to normalize each of the error matrices. The normalization allows error matrices with differing numbers of accuracy sites to be compared without calculating the Kappa statistic. One of the most prevalent processes used to normalize error matrices is called MARGFIT. This process uses iterative proportional fitting to balance each row and column so that they each sum to a specified number (Congalton et al., 1983; Congalton and Green, 2009). Often, the specified number is 1.0 so that the numbers in the error matrix can more easily be converted to percentages. Each cell in the normalized error matrix can then be viewed as a proportion of the specified number (Table 3). These proportions negate the need for user's and producer's accuracies because each of the cells along the major diagonal is the proportion of correctly identified accuracy sites. These proportions are representative of the probability for a classification site to fall in each cell, whether it be on the major diagonal or not, based on the given error matrix.

Table 3: A sample normalized error matrix (see Table 1 for original matrix).

Map Data	Agriculture	Barren	Forest	Lake	Row Total
	Reference Data				
Agriculture	0.7911	0.0572	0.0892	0.0622	1.00
Barren	0.1005	0.8357	0.0372	0.026	1.00
Forest	0.095	0.0851	0.7012	0.1189	1.00
Lake	0.0133	0.0221	0.1723	0.793	1.00
Column Total	1.00	1.00	1.00	1.00	4.00

Normalized Accuracy	0.7911+0.8357+0.7012+0.793=3.121	3.121/4=	**78%**

Since the normalization directly accounts for the error in the off diagonal cells, it may be the most inclusive accuracy assessment. However, it has been found that if an error matrix contains a significant number of zeros in the off diagonal cells, it may reduce the accurateness of the normalization. During the iterative process, the normalization will place positive numbers in the cells with zeroes and therefore change the meaning of the matrix, often with the result that the normalized accuracy does not agree with the overall or KHAT accuracies (Congalton, 1991). Many times this discrepancy can be resolved by using more accuracy sites.

Error in Accuracy Assessments

The error present in accuracy assessments can come from many different aspects of the classification process, from the data acquisition to the actual accuracy assessment process. One source of error in the accuracy assessment may involve the collection of the accuracy sites. By design, accuracy sites are often only collected in areas of homogeneous habitat, as discussed above. As suggested by Plourde and Congalton (2003, pg. 289), "Limiting the sampling to homogeneous areas of vegetation may inflate the accuracy measure for the map." Only taking accuracy sites in areas of homogeneous composition limits the acquisition to very specific sites, since most of nature is heterogeneous in makeup. The inflation of the accuracy occurs because often the areas on the map with homogeneous composition are also the easiest sites for the classification technique to identify. Therefore, a majority rule sampling of sites may be necessary. In a majority rule sampling technique the accuracy site is still taken within a certain size plot, but the plot no longer must be completely homogeneous in habitat type. The accuracy site is then assigned the reference data label of the habitat that makes up the majority of the sampling area. As a result, the accuracy sites are no longer limited to certain areas and are therefore more representative of the entire study area.

Although there is a clear advantage to using majority rule sampling sites, there is also a very important drawback. The majority rule sites can play havoc on the accuracy assessment if there are registration issues with the image. Since the pixels inside the accuracy site are mixed, a slight shift in the position of the site can put a whole new mix of pixels in the accuracy site possibly causing the majority habitat to change, making the reference data label incorrect (Plourde and Congalton, 2003). One suggestion for improving this problem is to have two rounds of sampling. The first round would be made up of homogenous accuracy sites and the second would be a random assortment of majority rule sites. This combination of the two sampling types may increase the success of the accuracy assessment by allowing sampling anywhere on the map, but also including some very clearly defined accuracy sites (Plourde and Congalton, 2003).

CHAPTER III

METHODS

Introduction

The methods used in this study were designed to test how each of the three hypotheses, outlined previously, may have contributed to the overall lack of improvement with editing in the benthic habitat maps created in Phase I of the Seagrass Monitoring Program. The first hypothesis suggests that some of the benthic habitats may be inseparable on the imagery creating some confusion during the classification and editing processes. The second hypothesis states that the sampling of the accuracy sites could have occurred most often in areas of homogeneous habitat, leading to an artificial inflation of the results of the accuracy assessment. The third and last hypothesis presumes that the small initial polygon size (i.e. segmentation scale parameter) of the Phase I imagery caused the lack of improvement with editing, since the image may have been over-segmented.

Therefore, this research explores how habitat type, sampling techniques, and initial polygon size might affect the appearance and accuracy of benthic habitat mapping. First, to test how these three attributes might affect habitat mapping, a statistical comparison of the benthic habitats that were most confused between the draft and contractor maps in Phase I was conducted. This analysis was done to assess whether the habitats were just too similar on the digital aerial imagery to accurately separate and identify. If so, it can be theorized that the Classification and Regression Tree (CART) technique used in this phase was insufficient and that additional methods must be used to differentiate these specific benthic habitats, or that no method will be able to differentiate these habitats. A Tukey's Honestly Significant Difference (HSD) test was employed as a way to compare the number of polygons in each group of confused benthic habitats with other groups of confused habitats in order to determine if one specific pair of benthic habitats were the hardest to separate and label on the imagery.

Secondly, the distribution of the accuracy assessment points in Phase I was studied to determine whether they were representative of the entire study area. Traditional sampling techniques suggest placing accuracy sites within homogeneous areas of habitat in order to avoid complications due to minor spatial shifts in accuracy sites (Congalton and Green, 2009). However, Plourde and Congalton (2003) found that taking accuracy sites exclusively in homogeneous habitats can cause artificial inflation of the accuracy assessment since these areas are usually also much easier to classify. The current study looks at the possibility that the accuracy sites may have been taken in homogeneous areas that were not representative of the diversity of the entire benthic area and therefore did not capture the area changed by the human photo-interpretation of the map. This analysis was completed by looking at the proportion of accuracy sites found in areas that are not traditionally used as accuracy sites, such as transitional habitats.

Lastly, a test was completed to determine whether the differences in the Phase I and II habitat maps may be contributed to the change in initial segmentation polygon size (Schiewe, 2002; Rahman et al., 2003). The initial draft map in Phase I had relatively small initial polygons, while the draft map for Phase II had larger initial polygons with a few smaller polygons in areas of mixed habitat. The resulting maps and accuracies of Phase I and Phase II were compared in this study to determine whether the smaller polygons actually reduced the effectiveness of the mapping technique. According to Schiewe (2002), the smallest minimum mapping unit may not always be the most effective.

The overall objective of this study was to complete a thorough analysis of how each of these techniques may have affected the creation and accuracy of the resulting benthic habitat maps. This objective was fueled by the observation that hundreds of hours of manual photo-interpretation vastly changed the benthic habitat map in Phase I without changing the overall accuracy of the map. Thus, this research was undertaken to investigate how this discrepancy may have occurred.

Reference Data

Phase I Data

Study Area

Recently, the State of Texas started a Seagrass Monitoring Program and began the exploration of their first study area with help from a number of scientists including Kass Green of Kass Green and Associates, Chad Lopez of Fugro EarthData International, and many others. Initially, there were six bays in the study area: Aransas; Copano; Redfish; Corpus Christi; Baffin; and Upper Laguna Madre Bays (Figure 5). The study site totaled over 1,400 square miles of coastline and included all of the underwater habitats of these bays as well as some of the surrounding low-lying land. The area also includes parts of two different watersheds; the Coastal Bend Watershed in the north and the Laguna Madre Watershed in the south. The bays are well protected from Gulf currents by barrier islands and the maximum depth for all of the bays is about 4.5 meters below sea level (not including man-made trenches) (Handley et al., 2007). The sheltered and shallow nature of these bays has allowed submerged seagrass meadows to become the dominant aquatic vegetation of the area, creating a perfect starting place for the Seagrass Monitoring Program (Green et al., 2007).

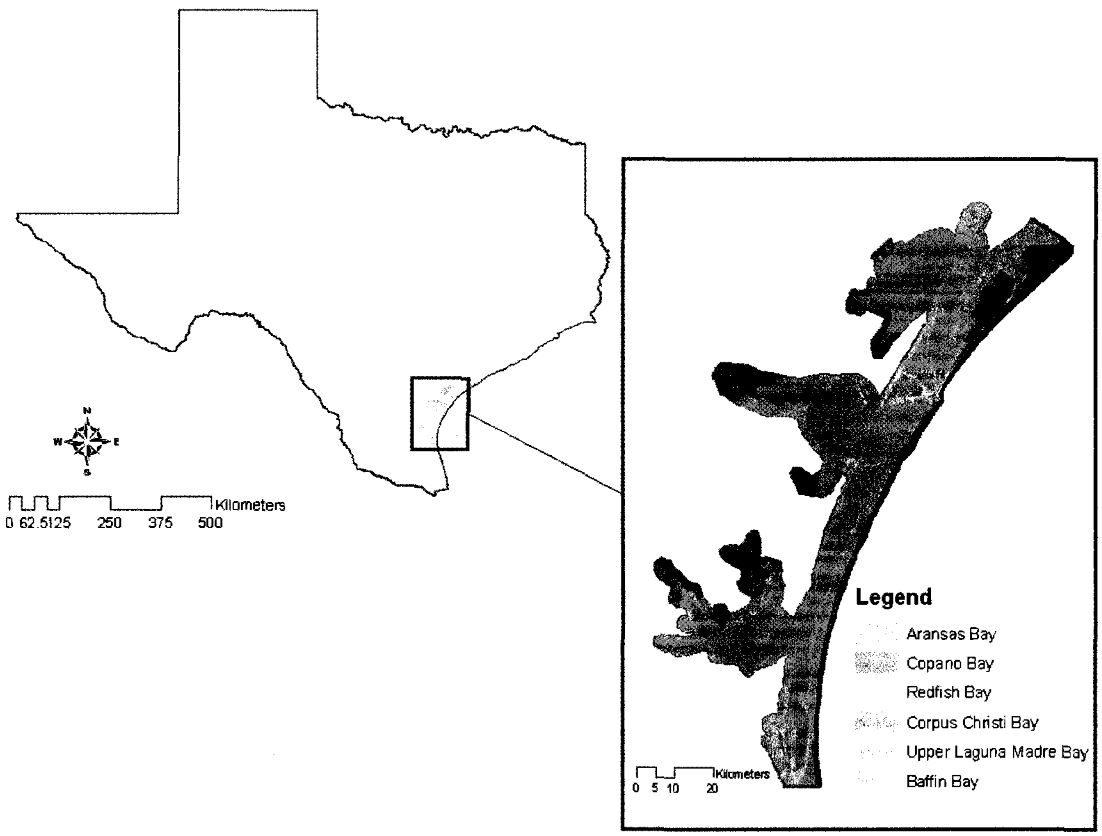


Figure 5: The entire Phase I study site in the southern Texas Gulf Coast.

Image Selection and Classification System

The images chosen for this study were compiled from a set of high-resolution digital aerial imagery taken for the National Agriculture Imagery Program (NAIP). This dataset was acquired in 2004 using a Leica ADS40 airborne digital camera with 12-bit, 1 meter pixel resolution. The imagery was then re-sampled to 8-bit, 2 meter pixel resolution for processing. The National Oceanic and Atmospheric Administration (NOAA) required that this project have a minimum mapping unit (mmu) of 0.01 hectares, which meant that 2 meter pixel resolution was more than sufficient for the purposes of these maps. This specific dataset was chosen for this project because, although it was captured for agricultural purposes, it was taken at a time of great water visibility and very little turbidity, making it perfect for benthic habitat mapping (Green et al., 2007).

In order to ensure that the classification system for these maps had labels and definitions that were mutually exclusive, hierarchical, and totally exhaustive, the scientists chose to create their system based off of the well established Florida System for Classification of Habitats in Estuarine and Marine Environment (SCHEME) (Appendix A) (Madley et al., 2002). SCHEME was originally intended for use in Florida, but was designed so that most general benthic habitats from around the globe were included and specific habitats could always be added or subtracted depending on later needs and the particular areas being studied. Therefore, a dichotomous key, which resulted in a set of rules, was created from SCHEME to be used specifically in this study area (Appendix B). The habitats encompassed in this key are displayed in Table 4.

Table 4: The habitats originally thought necessary for the Seagrass Mapping Project.

	Benthic Habitats
1	Continuous Macroalgae
2	Patchy Macroalgae
3	Continuous Submerged Rooted Vegetation (SRV)
4	Patchy SRV
5	Unconsolidated Sediments
6	Bivalve Reef
7	Tidal Swamp - Mangroves
8	Tidal Marsh - Spartina
9	Unknown Benthic Habitat
10	Hardbottom
11	Land
12	Mollusk Reef
13	Reef/Hardbottom

Pilot Study

An initial exploration of mapping methods was completed in Redfish Bay to determine the most effective and least costly method for mapping benthic habitats in the entire study area. Three classification techniques were tested in the pilot study and each of these techniques was completed using Feature Analyst by Visual Learning Systems. The three techniques included:

1) Classification using a Classification and Regression Tree (CART) technique;
2) Wall-to-Wall Classification; and
3) Feature-by-Feature Extraction.

All three of the classification techniques were applied to the same images using an OBIA approach (with the images segmented using Feature Analyst). Each technique also used the same field collected training and accuracy sites.

The classification techniques were conducted using the 11 habitats from the original dichotomous key (Table 4), but the list was then reduced to the actual habitats present

(Table 5). The accuracy for each of the maps was reported in error matrices and compared using the Kappa technique. Each of the maps resulted in accuracies in the 60%-65% range and they were not statistically different from each other (Green et al., 2007). Since the methods were not significantly different, any one of the three methods could be used with the same effectiveness as either of the others. Therefore, the best classification is the one that took the least amount of time and effort (Congalton and Green, 2009). The CART technique was subsequently determined to be the best technique for mapping because it was by far the fastest and least labor intensive (Green et al., 2007).

Table 5: The actual benthic habitats used in the pilot study.

	Pilot Study Habitats
1	Algae
2	Bivalve Reef
3	Continuous SRV
4	Patchy SRV
5	Land
6	Mangroves
7	Emergent Marsh
8	Unconsolidated Sediments
9	Unknown Benthic Habitat

After closely examining the resulting maps, the scientists determined that the Feature Analyst segmentation algorithm did not satisfactorily segment the image into polygons and may have contributed to the overall low accuracies. The team decided to try the same CART technique on polygons created by the Definiens Professional (i.e. eCognition) segmentation algorithm. This process resulted in a map that had acceptable polygon delineation and a classification that was not significantly different from any of the other methods performed (Green et al., 2007). Since this last method was the easiest to implement and resulted in the most visually pleasing polygons, it was chosen for use in the rest of Phase I of the benthic habitat mapping project.

Draft Map

Initially, a Definiens Professional segmentation algorithm with a scale parameter of 20 was applied to the imagery of the entire Phase I study area in order to get it ready for classification. The scale parameter of 20 produced over 2 million polygons with an average area of 0.17 hectares. In order to create a useful decision tree, a calibration trip into the field collected 583 field verified sample sites and a secondary 788 sample sites were recorded through manual photo-interpretation, producing 1329 total sample sites (Green et al., 2007). Another set of benthic habitat classes was identified and defined for the draft map to reflect the habitats found during the calibration trip (Table 6). A random number generator was then used to select and set aside 100 sample sites from each benthic habitat class to be used for accuracy assessment of the maps. The sample sites that were not selected for the accuracy assessment purposes were considered training sites and were used to define the rules for the CART technique that was implemented to create the draft map. These methods were largely the same as the ones used in the pilot study, only now including information gathered from the entire Phase I study area rather than just Redfish Bay.

Table 6: The draft map benthic habitat classes.

	Draft Habitats
1	Bivalve Reef
2	Continuous Macroalgae
3	Patchy SRV
4	Continuous SRV
5	Land
6	Mangroves
7	Spartina
8	Unconsolidated Sediments
9	Unknown Benthic Habitat

Once the rules were defined for the decision tree, each of the individual polygons was automatically classified using the same classification technique used in the pilot study, except it was based on the new rules. An accuracy assessment of the resulting map was made using the collected reference data. In the cases where the CART technique did not produce a useable map, some rules were slightly modified and the analysis was tried again. Once the highest accuracy was achieved, the rule set was defined as the final decision tree for the CART technique and the map was determined to be the draft map. Some slight manual editing was done to the draft map as it was produced, but overall it had very little human involvement and achieved an impressive deterministic accuracy of 83% (Green et al., 2007).

Contractor Map

Once the draft map was complete, a thorough manual edit of the map lasting over 500 hours was conducted, during which 26% of the polygons were changed (Green et al., 2007). First, per NOAA's request, all dock and bridge structures were taken out and the labels were replaced with the label from the benthic habitat underneath. A second major edit was that any large wetland or marsh area that did not contain any benthic habitats of interest, such as seagrass, was masked out of the classification and labeled as land for the assessment of the contractor map. Also in this step, since there were so few polygons with the label "Continuous Macroalgae", any polygons with this label were incorporated into the "Unconsolidated Sediments" category. Areas of known "Hardbottom" were also edited into the map. Therefore, with all of these modifications, the final classification scheme used to create this contractor map was different than that used by any of the previous maps (Table 7).

Table 7: The contractor map benthic habitat classes.

	Contractor Habitats
1	Bivalve Reef
2	Patchy SRV
3	Continuous SRV
4	Land
5	Mangroves
6	Spartina
7	Unconsolidated Sediments
8	Unknown Benthic Habitat
9	Hardbottom

Once most of the edits were complete, many of the smaller polygons were then dissolved into much larger polygons where the photo-interpreters felt it was appropriate. For instance, if several polygons all had the exact same label and shared an edge, they would be dissolved into one larger polygon. This process was to simplify the map as much as possible. The last edit to the contractor map was to smooth the polygons so they no longer reflected the boxy nature of the pixels of the original imagery. The deterministic accuracy of the contractor map subsequent to all of the editing and manual photo-interpretation was only 85%, not significantly different than the draft map which had very little human involvement (Green et al., 2007).

Since the reference data was used to check and modify the CART technique used to create the initial draft map, the data could no longer be assumed to be independent of the classification technique. Checking the CART technique with the reference data produced the best decision tree possible with the given data, however it may have caused the technique to be very good at classifying the reference data, but insufficient at classifying any other data. To verify that the accuracy sites were still fairly assessing the accuracy of the entire map, 287 more sample points were collected in a validation trip, and 74 of them were then randomly chosen to be used for another accuracy assessment. The remaining validation points that were not chosen to be used for the follow-up accuracy assessment were used to verify some specific points of confusion addressed in the editing process. Even with the new validation points and the independent reference data, the deterministic accuracy remained at 85% and the fuzzy accuracy was 90% for the contractor map (Green et al., 2007). Therefore, the original reference data retained their independence in the classification process, but the map was still not significantly better than the original draft map.

Final Map

A few more changes were made between the contractor map and the final benthic habitat map due to some suggestions made by NOAA. The habitat name "Spartina" was changed to "Emergent Marsh" so the habitat classes would more closely resemble the original dichotomous key habitat classes. In addition, the habitat name "Hardbottom" was changed to "Annelid Reef" to better portray the actual habitat present (Table 8). General editing was completed and the polygons on the Gulf of Mexico side of the barrier islands and the masked sections from the contractor map were removed from the map since they contained no habitats of interest for the project. Some "SRV" labels were pared down to be more conservative in their estimates and a "Mat Algae" modifier was added to the polygons that contained this habitat, despite it not being used in the classification. These minor changes in the map had no significant impact on the overall accuracy of the map. The deterministic accuracy of the final map was 86% and the fuzzy accuracy was 90% (Green et al., 2007).

Table 8: The final benthic habitat classes for the Phase I map.

	Final Habitats
1	Annelid Reef
2	Bivalve Reef
3	Patchy SRV
4	Continuous SRV
5	Land
6	Mangroves
7	Emergent Marsh
8	Unconsolidated Sediments
9	Unknown Benthic Habitat

Phase II Data

Study Area

The Seagrass Monitoring Program continued their study in 2008 focusing on two more bays along the seacoast of Texas; Lower Laguna Madre Bay and San Antonio Bay. Only the Lower Laguna Madre Bay was explored in this research study. Lower Laguna Madre Bay covers 800 square miles of coastline and is just south of the Upper Laguna

Madre Bay (Figure 6). Like the areas studied in Phase I, this bay is also shallow and protected with an average depth of 1 meter below sea level. The Lower Laguna Madre Bay is in the Laguna Madre Watershed and has the largest coverage of seagrass meadows of any bay in Texas, making it an important addition to the Seagrass Monitoring Program (Handley et al., 2007; Green et al., 2008).

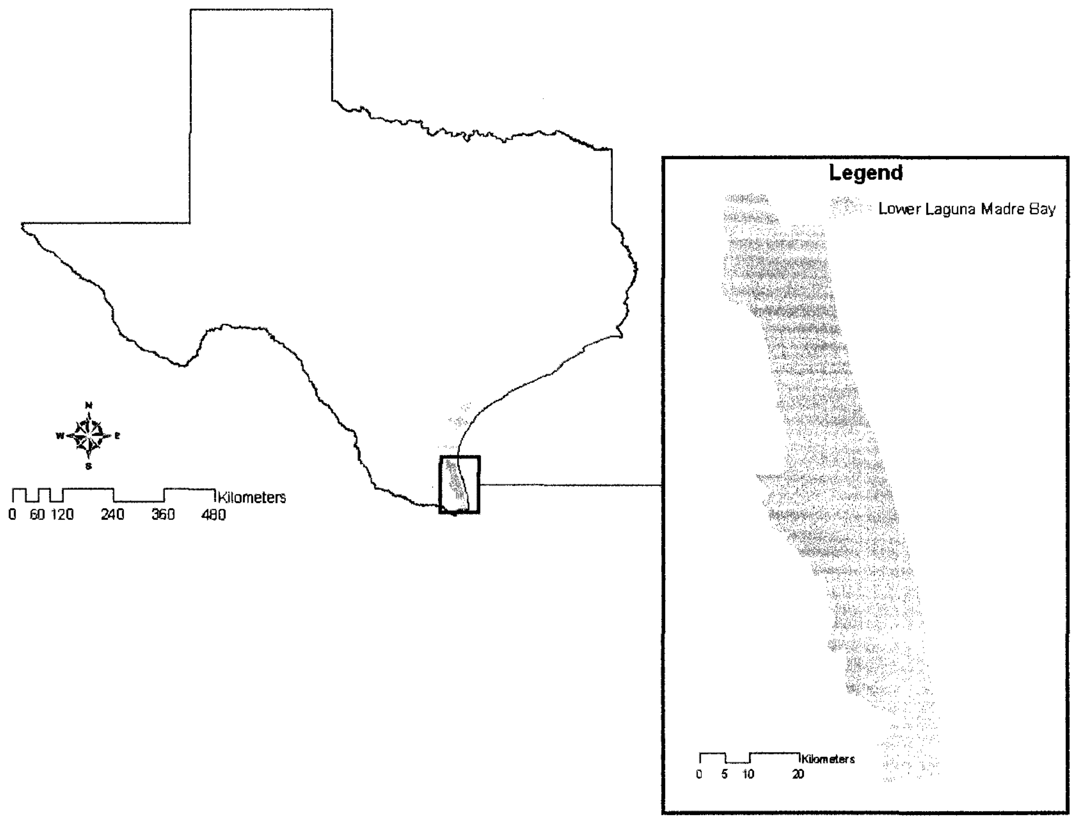


Figure 6: The Lower Laguna Madre Bay study site from Phase II.

Draft Map

As with the Phase I data, Phase II used the NAIP digital aerial images collected in 2004 as the imagery to be classified. The same Definiens Professional segmentation software was applied to the Phase II data, only this time the scale parameter was raised to

100 for the initial segmentation. The larger scale parameter led to an average polygon area of 3.12 hectares as opposed to the smaller 0.17 hectare Phase I polygons. With the larger initial polygons in Phase II, a different set of habitat labels were needed so that areas of mixed habitat could be identified (Table 9). As in Phase I, the same method of training the decision tree with both calibration sites and then correcting using the reference data, was used to classify the initial Phase II data. A total of 242 accuracy sites were collected, and 220 of them were used as the original reference data. The remaining 22 sites were kept independent of the CART technique and used to ensure that the accuracy assessments were fair (Green et al., 2008).

Table 9: Benthic habitat classes used in the classification of the polygons segmented with the scale parameter of 100.

	CART I Habitats
1	Continuous SRV Peppered
2	Continuous SRV Peppered
3	Emergent Marsh
4	Emergent Mix
5	Land Flats
6	Land Other Mix
7	Land
8	Mangrove Mix
9	Mangroves
10	Patchy SRV
11	SRV/Sediment Mix
12	Unconsolidated Sediments Algal
13	Unconsolidated Sediments
14	Unknown Benthic Habitat

After the first classification was completed on the 100 scale polygons, the mixed habitat polygons were then segmented into smaller polygons using the scale parameter of 20 in order to more precisely categorize the mixed habitats. Once the second segmentation was complete the average size of the polygons was 1.45 hectares and the benthic habitat classes more closely resembled those in Phase I, with the exception that no reefs were added at this stage and the "SRV/Sediment Mix" habitat label was kept for the time being (Table 10). Each mixed habitat was individually run through a CART technique with rules specifically designed to separate the mixed habitats in question. This classification process included very little manual editing and produced a map with a deterministic accuracy of 75% and a fuzzy accuracy of 76% (Green et al., 2008).

Table 10: The benthic habitat classes used after the mixed polygons were segmented into much smaller polygons and labeled individually.

	CART II Habitats
1	Continuous SRV
2	Emergent Marsh
3	Land
4	Mangroves
5	Patchy SRV
6	SRV/Sediment Mix
7	Unconsolidated Sediments
8	Unknown Benthic Habitat

Contractor Map

The Phase II editing did not take nearly as long as the Phase I editing since there were far fewer original polygons. Regardless, 29% of the polygons did change labels, only this time the accuracy improved with editing. As in Phase I, general edits were also present in Phase II, such as dissolving similar adjacent polygons together and smoothing of the polygons, as well as a slight change in the labels used (Table 11). Photo-interpretation was used to separate the few last remaining "SRV/Sediment Mix" polygons and known "Bivalve Reef" locations were edited into the map. Overall, the deterministic accuracy of the contractor map was 89% with a fuzzy accuracy of 90%, a significant improvement over the draft map (Green et al., 2008).

Table 11: The contractor benthic habitat classes for Phase II.

	Contractor Habitats
1	Bivalve Reef
2	Continuous SRV
3	Emergent Marsh
4	Land
5	Mangroves
6	Patchy SRV
7	Unconsolidated Sediments
8	Unknown Benthic Habitat

Final Map

The final map is based on suggestions that NOAA gave to the scientists following their own independent investigation of the area. Other than very minor changes, the final map very closely resembled the contractor map, with a deterministic accuracy of 90% and a fuzzy accuracy of 90%. As expected, these accuracies were not significantly different than those of the contractor map (Green et al., 2008).

Analysis

The overall goal of this study was to determine what might have caused the negligible change in accuracy observed after considerable editing of the benthic habitat map created in Phase I of this project. Three analyses were completed in this study to test the three proposed hypotheses. The first analysis was designed to determine whether there were habitats that were most often confused during the creation of the benthic habitat maps. The second analysis tested whether the placement of the accuracy sites was biased toward areas of homogeneous habitat, and the third analysis compared the results of the Phase I and Phase II maps to determine if the different segmentation parameters significantly changed the creation and accuracy of the two maps.

Hypothesis 1: Confused Habitats

Benthic habitats can be some of the most difficult habitats to map using remote sensing. In many instances it can be very difficult to separate benthic habitat classes by spectral reflectance because they are all under water. Therefore, it seemed quite plausible that some of the benthic habitats in the Phase I map could be spectrally confused, leading to many habitat misclassifications during the creation of the benthic habitat map. Imagery, GIS, and statistical analyses were all used in conjunction to test this theory.

Initially in ArcMap 9.2, a GIS, the draft map layer for each bay in Phase I was spatially joined with the matching contractor map layer so that the resulting polygons had the habitat labels from both the draft and the contractor maps. The spatial join was not perfect because it used both the boundaries of the draft map and the contractor map to define the boundaries of the polygons in the new layer. In some cases the contractor polygons had been smoothed to a point where they did not even closely resemble the original draft polygons, and this situation resulted in sliver polygons with no contractor label. In these instances, some manual editing was completed so that these slivers were given the label of the contractor polygon in which its centroid resided. Once the join was complete, the polygons that had differing labels between the draft and contractor habitat maps were selected and then turned into their own layer (Figure 7). These selected polygons represent the polygons that were changed or edited between the draft map and contractor map. This layer will later be referred to as the "changed layer" (i.e. the layer of changed polygons).

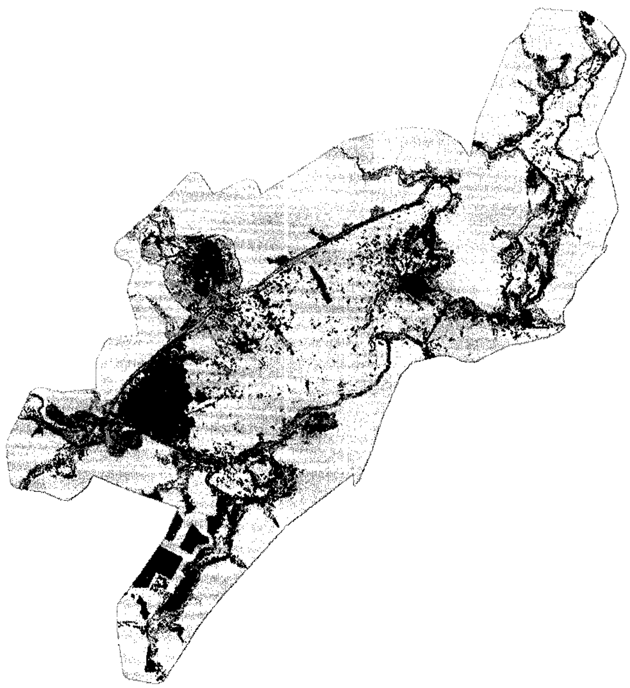


Figure 7: Copano Bay after selecting the polygons with different draft and contractor labels. The selected polygons are in red.

Each changed layer for each bay was then summarized by the number of times one specific draft habitat label was changed into another specific contractor habitat label. For instance, the number of times "Spartina" in the draft map changed to "Land" in the contractor map was counted and recorded. This summation was completed for each set of habitats and each bay. For simplicity, any "Continuous Macroalgae" sites were relabeled as "Unconsolidated Sediments" sites in the draft map since they had been modified to that label in the contractor map. Also, the "Hardbottom" sites in the contractor map were ignored since there were an insignificant number of them and had been manually added in the editing process, therefore not existing in the draft map.

To begin to determine whether one of these switched habitats was significantly more prevalent than any other, the changes for each bay were first summed. This sum included changes in either direction so that the two habitats that were most inseparable on the images could be determined without regard to the direction in which the labels were changed. For example, if 100 polygons total from one bay changed from the "Land" draft label to the "Spartina" contractor label and 50 polygons changed from "Spartina" to "Land", the two numbers where then summed to make 150 polygons total changed between "Spartina" and "Land" for that bay. These totals were then turned into percentages per bay so that if one bay had significantly more changed polygons than another, the differences in bays would be accounted for in the analysis. This analysis resulted in six different percentages (one for each bay) for each pair of habitats that had different labels in the draft map and the contractor map. Each pair of confused habitats is referred to as a group. If a group was not present in a bay, it was given a 0% for that bay to ensure that each group had six numbers. A Tukey's HSD post hoc test was completed using SPSS, a statistical software package, to determine whether any groups of changed habitats were significantly larger than any other group using the equation as follows:

$$q = \frac{M_1 - M_2}{\sqrt{MS_w\left(\frac{1}{n}\right)}}$$

where M is the group mean, MS_w is the mean square within the group, n is the number of percentages per group (always six in this case), and q is the Tukey's score that shows whether one group is significantly different than another.

The seven most confused habitat pairs were then broken down into two parts to account for the direction of change. Only the top seven habitat pairs were chosen because they were significantly larger than any of the remaining groups and encompassed the majority of changed habitats. These seven groups were split into their direction of change to show any planned editing in one direction that may have inflated the results or possible biases toward one label over another. For example, the new analysis would account for the change in direction from "Land" in the draft map to "Spartina" in the contractor map and vice versa by creating two separate groups. This new analysis resulted in 14 new groups with six percentages each. Another Tukey's HSD test was completed for these 14 groups to determine if one specific draft habitat more frequently was edited into another specific contractor habitat. A final Tukey's HSD test was used to determine whether the bays were significantly different enough to affect the number and type of confused pairs.

The reference data labels of the accuracy sites were then compared with the same site labels from the changed habitat layer to determine the number of accuracy sites that changed from right to wrong, wrong to right, or wrong to wrong. Since the accuracy sites only had one correct reference data label, no sites were changed from right to right. The percentage of each type of changed accuracy site was calculated for each group of confused habitats. For instance, if 24 accuracy sites were changed from right to wrong going from "Spartina" to "Land", it may only account for 7% of the total polygons that were changed from right to wrong. However, if 10 labels went from wrong to right changing from "Unconsolidated Sediments" to "Patchy SRV", that may account for 8% of the wrong to right sites. A MARGFIT analysis was completed on the contractor map error matrix to determine if the most confused accuracy sites on the contractor map matched the highest number of wrongly changed polygons. The MARGFIT analysis was designed so that all rows and columns summed to 1.00, so each normalized value could be easily represented as a percent chance that an accuracy site would fall into that specific box.

Lastly, similar tactics were used on the Lower Laguna Madre Bay in Phase II to determine whether any group of habitats was significantly more changed than any other group. A MARGFIT analysis was also used to determine the most misclassified habitat. These results were then be compared to the Phase I results using a Kappa analysis to determine whether one classification was significantly different from the other.

Hypothesis 2: Accuracy Site Collection

Traditionally, reference data are usually collected in areas of homogeneous habitat (Plourde and Congalton, 2003), but with benthic habitats, homogeneity is usually a characteristic of a vast minority of the areas being mapped. Therefore, this study set out to see if during the collection of the reference data, the scientists had avoided areas of transitional habitat and therefore had artificially inflated the accuracy of the draft and contractor maps in Phase I (Plourde and Congalton, 2003). If the accuracy sites are in fact not in transitional habitats, it would explain why changing many of the habitats had no impact on the accuracy assessment, since the accuracy sites were only sampling areas that were easily classified by both the CART classification technique and the photo-interpretation methods used during editing.

In areas such as Texas, the controlling factor regarding where seagrasses, or "SRV", will no longer grow is determined by whether or not light can filter through the water to the sea floor (Onuf, 1996; Koch, 2001; Zajac et al., 2003; Handley et al., 2007; Lee et al., 2007). Two very important factors in this calculation are water clarity and water depth. In Phase I, areas of no growth can be approximated by a single water depth since the bays all have similar water clarities. According to Onuf (1996), the depths to which seagrass can grow in these bays ranges anywhere from 1.4 meters to 1.8 meters. However, Pulich states in the Seagrass Status and Trends in the Northern Gulf of Mexico USGS Scientific Investigations Report (Handley et al., 2007) that they do not usually see seagrass at depths greater than 1.5 meters. Therefore, a depth of 1.5 meters below sea level was chosen as the maximum depth of continuous seagrass growth in these bays.

For this study, using the depth of 1.5 meters as the midpoint, two transition zones were defined to try and capture areas that would have sparse seagrass transitioning to other habitats and therefore would not be optimal for the placement of accuracy sites. To accomplish this task, an ASCII format bathymetry layer of the area was retrieved from NOAA's National Geophysical Data Center (NGDC) Design-A-Grid site using the lower latitude of 26° 46' N, an upper latitude of 28° 21' N, a left longitude of 98° 18' W, and a right longitude of 96° 23' W. This data set was derived from the NGDC Coastal Relief Model with less than 90 meter horizontal resolution and 1/10 meter vertical resolution (Divins and Metzger, 2008). The ASCII file was then loaded into ArcMap, projected, and changed to a usable raster format. The raster was then converted to vector format so that 3D Analyst could produce 0.25 meter contours of the entire bathymetry dataset. The contour layer was then cut to the area of the Phase I study site and two transition areas were created using the geodatabase tools in ArcMap to convert the contour lines to polygons.

The first transition zone was defined as 1.25 meters to 1.75 meters below sea level and the second was a more conservative definition of a transition zone with depths from 1 meter to 2 meters below sea level (Figure 8). Using depth to define the transition zones was also helpful in that it accounted for the gradient of the sea floor. Since seagrass prefers a slowly changing gradient to a steep gradient, steeper slopes cause a more abrupt meadow edge, shortening the transition zone (Koch, 2001). The third and final transition zone served to account for the shallow water edge of seagrass meadows. Seagrass meadows normally do not begin until at least a depth of 0.5 meters because wave dynamics and tidal flow cause water that is too turbulent for segrass to effectively put down roots (Koch, 2001). Therefore, the last transition zone was found by selecting any draft polygons within 0.1 meters horizontal distance of the 0.5 meters below sea level contour line (Figure 9).

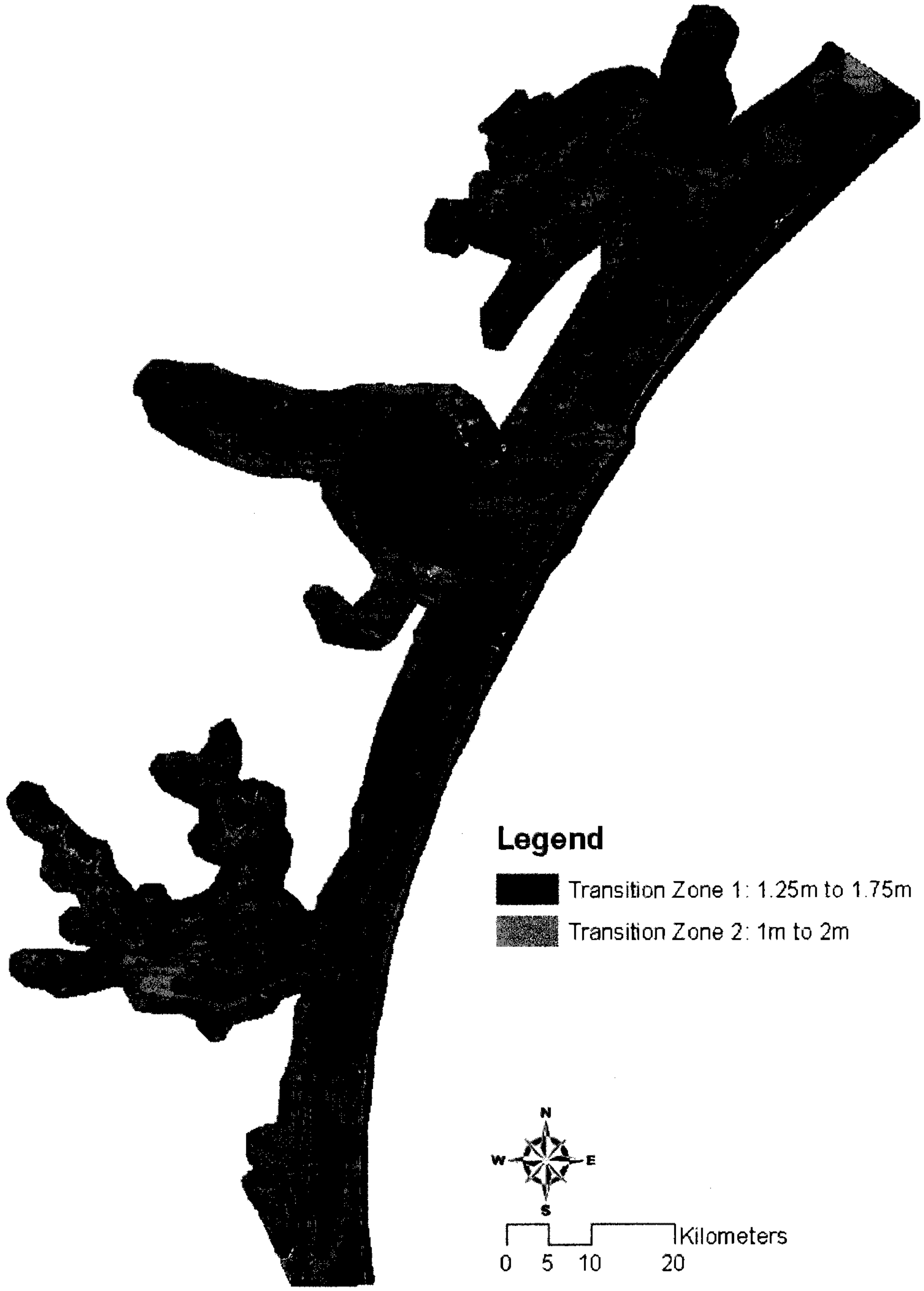


Figure 8: Transition zones 1 and 2 on the 2004 NAIP digital aerial imagery.

Figure 9: Transition zone 3 on the 2004 NAIP digital aerial imagery.

After the definition of these transition zones, the fraction of the total area that each transition zone occupies was calculated. This fraction was then compared to the percentage of the total accuracy sites centered in each transition zone to determine whether the accuracy sites were adequately sampling these areas. Lastly, the total area of changed polygons in each transition zone was calculated by selecting and summing the area of the polygons from the changed polygon layer, created during the exploration of the first hypothesis, which intersected the polygons from each transition zone. The results of this test were then used to determine whether the transition zones caused more confusion during classification than the non-transition zones.

Hypothesis 3: Polygon Size

The initial segmentation of the images can have a very important effect on the final outcome of the benthic habitat map and many times the smallest segmentation is not the most accurate (Schiewe, 2002; Rahman et al., 2003). To determine whether the initial scale parameter of 20 may have created polygons that were too small to produce the best benthic habitat mapping outcome, the draft map of Phase I was compared to the draft map of Phase II. The initial step necessary to compare the two maps was to dissolve the polygons in each map by habitat so that the true sizes of the habitats were represented, rather than the size of the segmentation polygons. Since the editing was completed on the dissolved draft map, dissolving the polygons also created a map that was more representative of the draft map used during the editing process. Once the habitats were dissolved, the average polygon sizes were computed for each map.

The next step was to determine whether the confused habitats from Phase I were significantly smaller than the average draft polygon size. This analysis might show that, because the initial polygon sizes were so small, the segmentation may have cut continuous habitats apart due to small variations in spectral reflectance. The variations may have been due to a shadow or surface light scattering and caused the classification technique to misclassify one small polygon within a larger group of polygons all within the same habitat (Schiewe, 2002). These smaller misclassified polygons should have been clumped into the larger habitat class during the map editing, thus creating a changed polygon without this change being detected by the accuracy assessment. The same analysis was done on the Phase II data to determine whether the changed polygons were also smaller than the overall average draft polygon size. However, in just the Phase I data, the size of polygons that were changed from wrong to right, right to wrong, and wrong to wrong were all compared to see if polygon size may have impacted the success of the editing process.

Lastly, the overall effectiveness of each of the classification algorithms was compared to see if different scale parameters changed how the final map was created. The time and effort to edit each of the maps was taken into consideration, as well as the number and percentage of corrections that were made and the accuracy of the different draft and contractor maps. Therefore, the benthic habitat map with the best results was determined not only by the overall accuracy of the map, but also by the speed and efficiency of the classification and editing processes applied to the map.

CHAPTER IV

RESULTS

The overall objective of this study was to determine how the accuracy of the benthic habitat map for Phase I of the Seagrass Monitoring Program did not improve between the draft and the contractor maps, even with extensive manual editing. The study was broken into three sections to test each of the three hypotheses. The first test was to see if there was a pair of habitats that was most confused, or changed, between the draft and contractor map to determine if these habitats were too similar to separate on the imagery. Secondly, the placement of the accuracy sites was tested to determine whether they are statistically not in areas of transitional habitat. The last test was to determine whether the segmentation in Phase I over-segmented the image making the classification much more difficult to complete, and to compare the resulting benthic habitat map of Phase I with the Phase II map created using a larger segmentation scale parameter.

Hypothesis 1: Confused Habitats

During the manual editing portion of Phase I of this project, a total of 429,428 (or 26%) of the original draft polygons changed labels during the editing process to make the contractor map. These changed polygons were then categorized into groups according to the two labels the polygon had been given for the draft and contractor maps. For instance, a group might contain all polygons that were labeled "Unknown Benthic Habitat" in one map and "Mangroves" in the other. Of all of the groups created, the "Spartina/Land" group claimed the highest percentage of changed polygons with 25.78% (Figure 10). According to a Tukey's Honestly Significant Difference (HSD) test performed in SPSS, the mean percent change of all of the bays collectively for the "Spartina/Land" group is significantly larger than any other group mean at a significance level of 0.05. The next subset of groups with similar means was: "Unconsolidated Sediments/Land", "Spartina/Unconsolidated Sediments", "Bivalve Reef/Unknown Benthic Habitats", "Unconsolidated Sediments/Continuous SRV", "Patchy SRV/Continuous SRV", and "Unconsolidated Sediments/Patchy SRV". Within this subset, the groups do not have means that are significantly different from each other, but the means are significantly greater than the remainder of the groups.

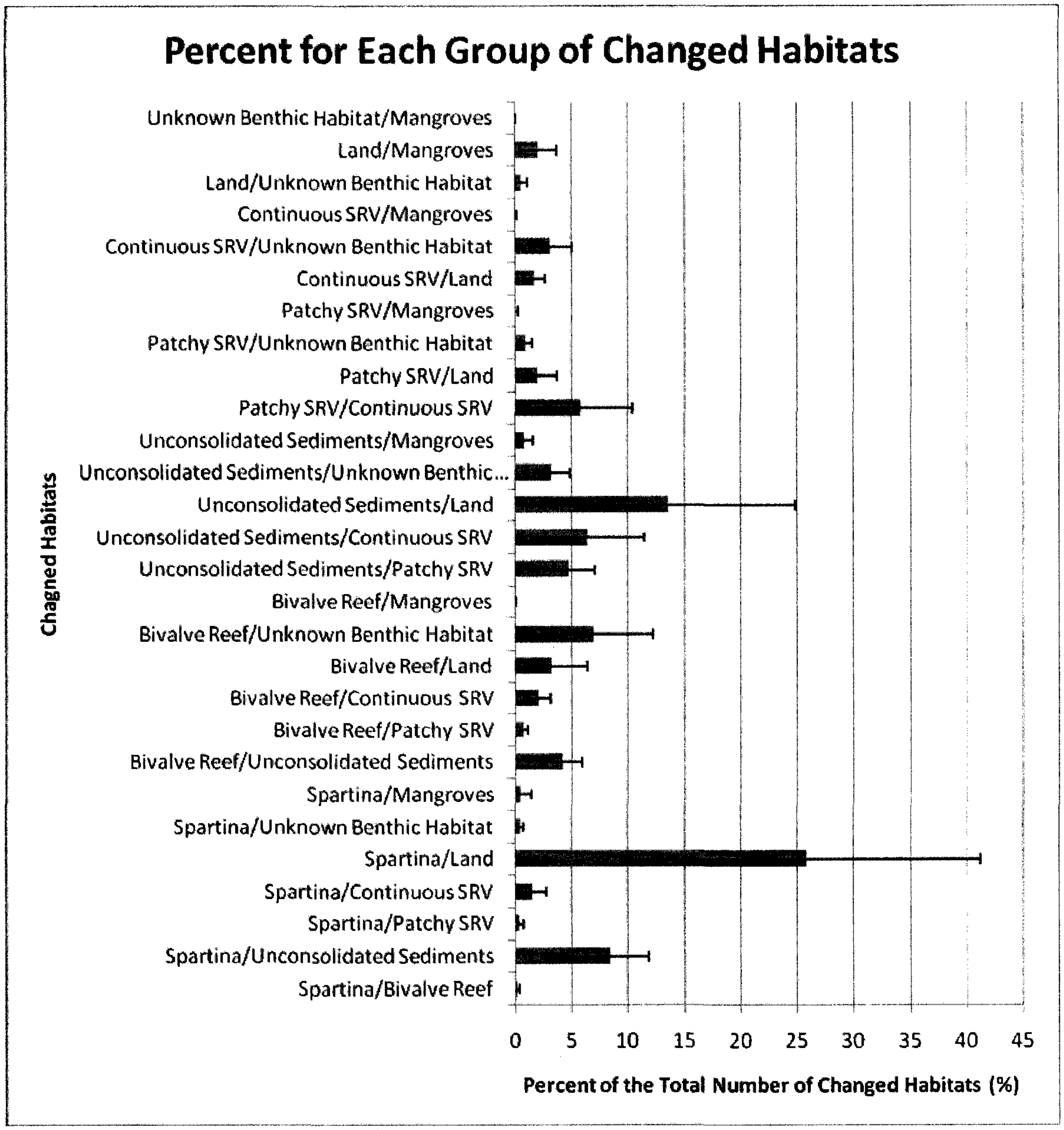


Figure 10: The percentage of the total number of changed habitats for each group with standard deviations displayed.

Because the top seven changed habitat groups had means that were significantly larger than the remainder of the changed group means, these seven were chosen for further investigation. First, the top seven groups were split depending on the direction of the change, therefore resulting in 14 new groups. These groups have labels that depict exactly which label the polygons initially were given in the draft map as well as the label the editing changed the polygons to in the contractor map. For example, the "Spartina/Land" group was split into two groups: the "Spartina to Land" group and the "Land to Spartina" group. Another Tukey's HSD test determined that the "Spartina to Land" group had far and away the highest number of changed polygons with 25.16%, which is significantly larger than any other group at the 0.05 level (Figure 11). The remaining 13 groups were not significantly different from each other. A final Tukey's HSD test determined that the six bays were not significantly different from each other in their number and type of habitats at the 0.05 significance level.

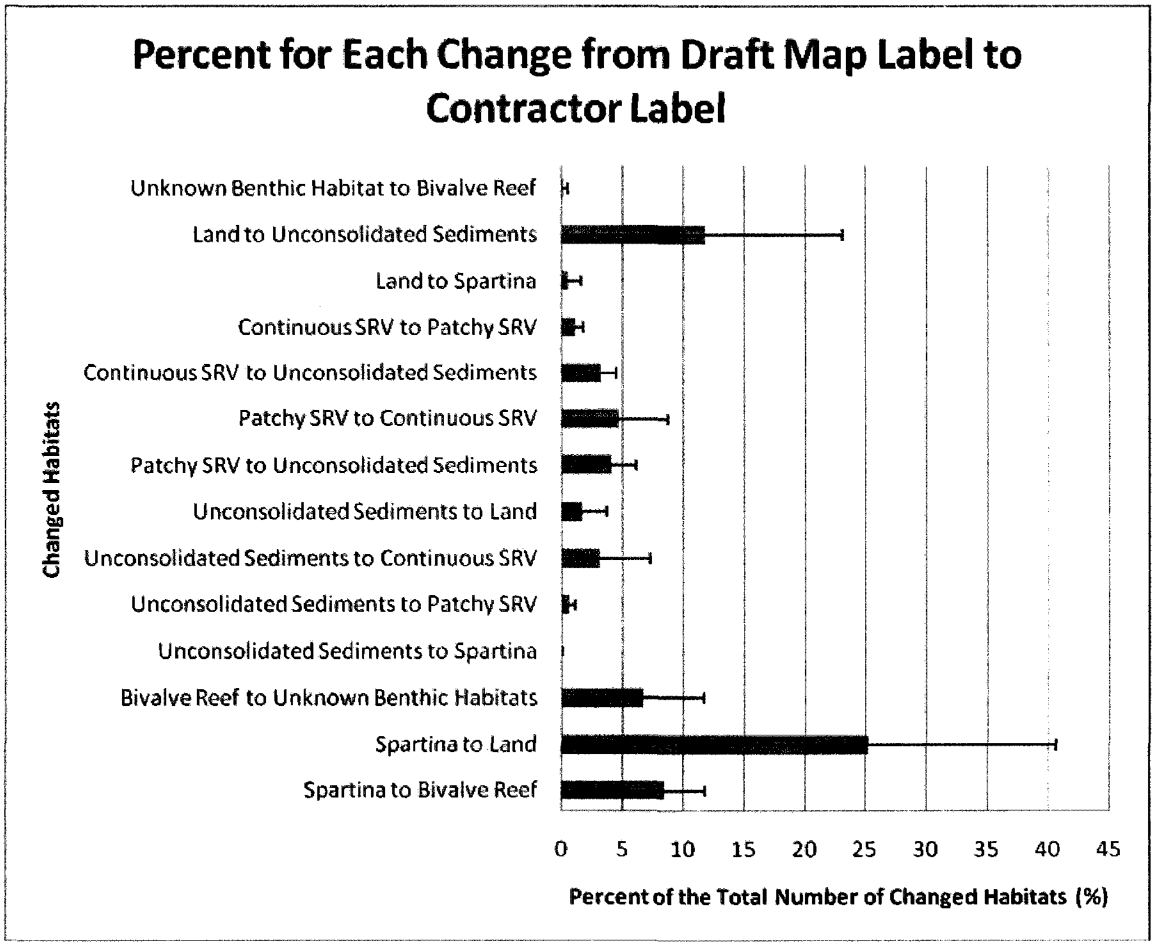


Figure 11: The top seven groups of changed habitats split into the direction of change (draft to contractor label).

In order to more specifically look at how sites were changed from the draft map to the contractor map, the reference data points (i.e. accuracy sites) that coincided with polygons within the changed layer were split into groups according to their draft and contractor labels. In total, 252, or 32%, of the total number of accuracy sites coincided with changed polygons. These polygons were then categorized by whether their labels changed from wrong to right, right to wrong, or wrong to wrong. Of the changed polygons that coincide with accuracy sites, there are 116 polygons that changed from wrong to right, 69 from right to wrong, and 37 from wrong to wrong. Since fuzzy accuracy methods were not included in this portion of the study, no polygons were changed from right to right.

In the wrong to right category, polygons that were originally labeled as "Continuous SRV", and then edited to the correct label of "Bivalve Reef", make up the largest group and account for 9.48% of the wrong to right polygons. The next two largest groups in this category are the "Unconsolidated Sediments to Continuous SRV" group and the "Continuous SRV to Unconsolidated Sediments" group at 8.62% and 6.90%, respectively.

By far the largest group of polygons that were edited from right to wrong is the group that started with a draft label of "Spartina" and was changed into "Land". This group accounts for 34.78% of the right to wrong polygons. The second and third largest groups in the right to wrong category are the "Unconsolidated Sediments to Continuous SRV" and the "Spartina to Unconsolidated Sediments" groups, with 11.59% and 10.14% of the category, respectively.

The wrong to wrong category is slightly more complicated because in this category, a polygon's draft and contractor labels are two different labels, and the coinciding reference data point label is something else entirely. Therefore there are three different labels for each polygon. 24.32% of the polygons in the wrong to wrong category have a draft label of "Continuous SRV", a contractor label of "Patchy SRV", a reference data label of "Unconsolidated Sediments". The second and third largest groups comprise 10.81% of the polygons in the wrong to wrong category each. Both groups also have a contractor label of "Continuous SRV", and a reference data label of "Unconsolidated Sediments", although one group has a draft label of "Bivalve Reef" and the other has a draft label of "Patchy SRV".

In order to compare the changed polygons with the accuracy of the contractor map, a normalized contractor map error matrix was created to determine which labels had the most misclassifications. The normalization was completed using MARGFIT with convergence after 26 iterations and a maximum deviation of <0.001. In the normalized contractor map error matrix, it is evident that the most confused labels are classified as "SRV" on the contractor map and have a reference data label of "Unconsolidated Sediments" (Table 12).

Table 12: Normalized contractor map error matrix for Phase I produced using MARGFIT. Each row and column sum to one and each cell was converted to a percentage of one. The major diagonal is in gray and each box represents the percent probability for an accuracy site to be in that category based on actual accuracy site results. Spartina is labeled as Emergent Marsh to match the final habitat error matrix.

	Reference Data						
Map Data	Bivalve Reef	SRV	Land	Mangroves	Emergent Marsh	Unconsolidated Sediments	Unknown Benthic Habitat
Bivalve Reef	91.05%	1.62%	1.08%	0.58%	0.34%	4.05%	1.28%
SRV	1.42%	86.17%	0.29%	0.16%	1.54%	10.07%	0.35%
Land	0.19%	0.64%	90.91%	1.61%	5.70%	0.44%	0.51%
Mangroves	0.40%	1.37%	0.91%	95.55%	0.28%	0.31%	1.08%
Emergent Marsh	1.08%	3.68%	2.45%	1.32%	87.71%	0.84%	2.92%
Unconsolidated Sediments	3.29%	4.79%	3.20%	0.57%	4.31%	82.58%	1.27%
Unknown Benthic Habitat	2.56%	1.74%	1.16%	0.21%	0.12%	1.72%	92.59%

In Phase II, the group with the highest number of total changed polygons is the group with the "Unconsolidated Sediments" label in the draft map and the "Land" label in the contractor map. This group contains 18% of the 21,644 changed polygons. The next two largest groups are the groups that were relabeled from "Unconsolidated Sediments" to "Emergent Marsh" with 16% of the total changed polygons, and "Unconsolidated Sediments" to "Continuous SRV" with 9%. Again, MARGFIT was used to normalize the contractor map to determine whether the changed polygon count matched the most common error in the contractor error matrix (Table 13). This normalization was completed in 16 iterations and a maximum deviation of <0.001. Unlike the contractor error matrix in Phase I, the highest probability to be misclassified is found in the group that is labeled as "Unconsolidated Sediments" on the contractor map and has a reference data label of "SRV". Despite this difference, the contractor map for the Phase I data and the contractor map for the Phase II data are both significantly better than a random classification (KHAT≥0.823), but are not significantly different from each other (Z=1.478).

Table 13: Normalized contractor map error matrix for Phase II produced using MARGFIT. Each row and column sum to one and each cell was converted to a percentage of one.

	Reference Data						
Map Data	Bivalve Reef	SRV	Land	Mangroves	Emergent Marsh	Unconsolidated Sediments	Unknown Benthic Habitat
Bivalve Reef	87.96%	1.65%	2.06%	1.75%	1.37%	2.46%	2.83%
SRV	1.75%	80.51%	1.27%	3.24%	0.84%	10.62%	1.75%
Land	2.05%	1.19%	90.71%	1.26%	0.99%	1.78%	2.05%
Mangroves	2.07%	1.20%	1.50%	80.34%	10.96%	1.79%	2.06%
Emergent Marsh	3.05%	1.78%	2.21%	9.40%	77.82%	2.65%	3.04%
Unconsolidated Sediments	1.69%	12.83%	1.23%	3.13%	7.35%	72.05%	1.69%
Unknown Benthic Habitat	1.42%	0.83%	1.03%	0.88%	0.69%	8.64%	86.58%

The objective of hypothesis one was to determine whether a set of benthic habitats were spectrally or visually inseparable on the digital imagery and caused the large change in polygon labels observed in Phase I of this project. Despite the fact that the most confused habitats were "Unconsolidated Sediments" and "SRV", as seen in the normalized error matrix, no one group of changed habitats was significantly larger than any other groups of changed habitats in Phase I. Therefore, no two "inseparable" habitats could have caused the large change in polygon labels during editing and the null hypothesis for hypothesis one cannot be rejected.

Hypothesis 2: Accuracy Site Collection

The first transition zone studied, from 1.25 to 1.75 meters below sea level, contains 14,047.05 hectares of habitat, or 4% of the entire Phase I study area. This zone also has 71 accuracy sites, or 9% of the 785 total accuracy sites collected. Transition zone two encompasses all habitats between 1 and 2 meters below sea level. This zone has 8% of the total area of the study site with 32,551.25 hectares of habitat. Transition zone two also has 14% of the total accuracy sites, numbering 112 sites. Lastly, transition zone three includes any habitats within 0.1 meters of the 0.5 meter below sea level contour line. This zone encompasses 2% of the study site area with 9,821.67 hectares of habitat and has 20% of the total accuracy sites with 159 sites (Table 14).

The changed polygons described in Hypothesis 1 were also investigated with regard to these three different transition zones (Table 14). Eight percent of all of the changed polygons are contained within transition zone one, while 12% are contained within transition zone two and 13% are contained within transition zone three. These percentages amount to 5,726.60 hectares of changed polygons in zone one, meaning 41% of zone one is made up of changed polygons. Similarly, in zone two 27%, or 8,827.40 hectares, of the transition area is made up of changed polygons. Therefore, transition zone three, although the smallest in size, had the largest area of changed polygons at 9,708.48 hectares, meaning 99% of this zone is made up of changed polygons.

Table 14: The proportion of habitat, changed polygons, and accuracy sites in each transition zone.

Transition Zone Proportions

Transition Zone 1:	1.25m depth to 1.75m depth
Transition Zone 2:	1m depth to 2m depth
Transition Zone 3:	Within 0.1m of 0.5m depth

Habitat	Area (hectares)	Percent of Total Study Site Area
In transition 1	14047.05094	4%
In transition 2	32551.25211	8%
In transition 3	9821.665985	2%
Total Area in Study Site	399840.8225	

Changed Sites	Area (hectares)	Percent of Total Changed Area	Percent of Transition Zone
In transition 1	5726.599199	8%	41%
In transition 2	8827.400798	12%	27%
In transition 3	9708.480792	13%	99%
Total Changed Area	73111.06264		

Accuracy Sites	Number of AA Sites	Percent of Total AA Sites
In transition 1	71	9%
In transition 2	112	14%
In transition 3	159	20%
In changed area	252	32%
Total AA Sites	785	

Hypothesis two was used to test whether the accuracy sites visited by the researchers were collected in specific areas to avoid transitional habitat. The results show that the proportion of accuracy sites within each transition zone is actually larger than the proportion of area covered by that transition zone, meaning that the transition zones were not under-sampled. However, a χ^2 test between the expected proportion of accuracy sites in each transition zone (equal to the percent area covered by that transition zone) and the observed proportion of accuracy sites (Table 14) showed that the two numbers were significantly different ($p<0.05$). Therefore, the researchers actually over-sampled each of the transition zones, possibly decreasing the results of the accuracy test rather than inflating the results. Therefore, the null hypothesis for hypothesis two cannot be rejected.

Hypothesis 3: Polygon Size

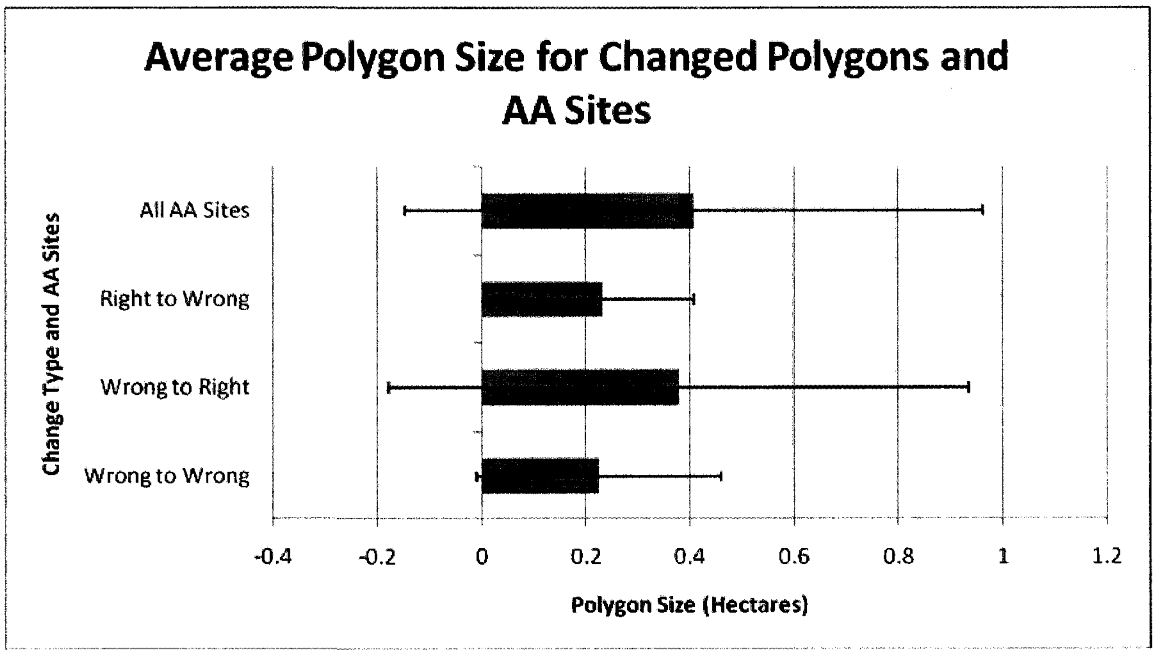
In the original draft map for Phase I, the segmentation process delineated over 2 million polygons, but by the time the contractor map was created, there were only 51,770 polygons. The average polygon size in the original draft map segmentation is 0.172 hectares, since a minimum mapping unit of 0.01 hectares was enforced. The average polygon size in the contractor map is around 8.063 hectares, which is significantly larger than the average polygon size in the original draft map. This difference in polygon size may be even more substantial because the distribution of the polygons was such that most of the polygons in the original draft map were smaller than the reported average with a few larger polygons (Appendix C).

As previously discussed, during the initial editing of the draft map, one of the first processes performed was to dissolve the smaller segmented polygons into larger groups of the same label. This process may account for a large part of the discrepancy in polygon size between the draft and contractor maps, since the dissolved average polygon size is more representative of the map that the researchers used for the manual editing process. The total number of polygons in the draft map after the dissolving process is 113,682, still more than twice the contractor map, and the average polygon size is 4.289 hectares. Finally, the average size of the 99,327 polygons that were changed, or edited, between the dissolved draft map and contractor map is 0.736 hectares.

The same procedure was completed for Phase II of the study with fairly different results. The original draft map contains 66,353 polygons with an average area of 3.117 hectares. When dissolved, the draft map has a larger average polygon size of 9.548 hectares and contains 21,664 polygons. The contractor map has larger polygons still, with an average polygon size of 10.589 hectares, as well as only containing 19,534 polygons. In Phase II, there are 21,531 confused polygons between the draft and contractor maps and they have an average size of 2.252 hectares (Appendix C).

The polygon categories defined in the first hypothesis by type of change (i.e. wrong to right, right to wrong, or wrong to wrong) were also analyzed for their average polygon size. As displayed in Figure 12, the average polygon size was largest for the group of polygons that changed from wrong to right and this average was very similar to the overall average accuracy site size. The two groups of changed polygons that ended in a wrong classification were similar to one another in average and had the smallest average polygon sizes.

Figure 12: The average polygon size for changed accuracy site polygons and the average size of accuracy assessment (AA) sites. Averages are in blue and the bars represent the standard deviation.



Hypothesis three was used to determine whether the initial segmentation scale parameter implemented in Phase I over-segmented the image, causing the mapping techniques to be less effective than those used in Phase II. In both phases of the project the majority of the changed polygons were much smaller than the average polygon size, meaning most of the editing was completed on the smaller polygons of the maps. However, the polygons that were edited incorrectly were even smaller than the average changed polygon size, indicating that the larger polygons were easier to edit and the smaller polygons in Phase I were much harder to correctly edit when compared to the larger polygons in Phase II. Therefore, the null hypothesis for hypothesis three can be rejected.

CHAPTER V

DISCUSSION

In this research, analyses were completed to determine what may have caused the lack of improvement with editing of the draft map for Phase I. The first analysis showed that there were a number of habitats that were changed or relabeled between the initial draft map and the edited contractor map. The second analysis defined three different transition zones and determined the number of accuracy sites within each of these transition zones. The third and final analysis used information from the first two studies to determine how the initial polygon size used in the draft maps may have affected the accuracy of the final maps in Phase I and Phase II.

Hypothesis 1: Confused Habitats

Of all of the changed polygon groups studied, the "Spartina to Land" group was significantly the largest group of polygons that were changed between the draft map and the contractor map. However, as discussed previously, many of the labeled "Spartina" areas were masked into "Land" for the contractor map (before being removed in the final map) since they did not contain benthic habitats of any concern to the group mapping the area. Therefore, this large group of changed habitats is expected. Similarly, this masking of the "Spartina" polygons accounts for the large number of polygons changed from right to wrong when changed from “Spartina” to “Land”, since the reference data labels of the accuracy sites were not changed to reflect the masking. No other change from one habitat to another is significantly more prevalent than any other change, meaning there are no two habitats that were significantly the most confused during the manual editing process.

During the creation of an error matrix for the contractor map, the reference data labels of the accuracy sites in the masked “Spartina” sections were changed to “Land” to account for the masking. In addition, the “Continuous SRV” and “Patchy SRV” polygons were combined into one large “SRV” category since the objective of the map was to find any “SRV” sites. As shown by the normalized contractor map for Phase I, the largest number of reference data points that were wrongly classified all have a classification label of “SRV” (Patchy or Continuous) and a reference data label of “Unconsolidated Sediments”. However, the error of commission for “Unconsolidated Sediments” polygons wrongly labeled as “SRV” may be a function of the design of the classification technique. The researchers who defined the decision tree that was used to classify the map were looking for areas of “SRV” and therefore would rather overestimate the existence of “SRV” than underestimate. It is much better for the Seagrass Monitoring Program to have to check areas for conservation on the ground that might have “SRV” and find none, than to completely miss viable areas of “SRV” habitat and therefore not care for these areas.

For Phase I, if the number of polygons changed between either of the “SRV” categories and “Unconsolidated Sediments” are combined, meaning “Patchy SRV” and “Continuous SRV” are treated as one category as done with the error matrix, the percentage of the total number of changed polygons contained in this group increases. The new "SRV/Unconsolidated Sediments" group accounts for 11.91% of the total changed polygons, becoming the third largest group of changed polygons. If the "Spartina/Land" group is taken out of the mix, since it is falsely large through masking, the "SRV/Unconsolidated Sediments" group represents the second largest group of changed habitats. However, even at 11.91%, this group is still not significantly larger than the other changed polygon groups.

Whereas the polygons labeled "SRV" on the contractor map with a reference data label of "Unconsolidated Sediments" are the most misclassified polygons in Phase I, the opposite is true for Phase II. In Phase II, the highest number of misclassifications occurs with the polygons classified as "Unconsolidated Sediments" on the contractor map and that have a reference data label of "SRV". This switch in categories clearly illustrates that while the greatest number of changed polygons may not have been between "SRV" and "Unconsolidated Sediments", the largest number of misclassifications is between these categories for both phases of the project. However, unlike in Phase I, many polygons with "SRV" reference data labels are misclassified as "Unconsolidated Sediments" in Phase II. Therefore, in Phase II, many areas that might make good conservation sites will be overlooked due to misclassification.

Since the "SRV/Unconsolidated Sediments" polygons are the most misclassified polygons and because they did largely contribute to vast number of changed polygons, the classification process or imagery techniques may not yet be refined enough to accurately separate and classify this pair of benthic habitats. Overall, these two habitats generally reside in the same areas, and "SRV" is usually found interspersed with

"Unconsolidated Sediments". Therefore, it can be quite difficult to determine where a small and thinly distributed patch of "SRV" exists when it is growing in a vast expanse of "Unconsolidated Sediments".

This difficult delineation between "SRV" and "Unconsolidated Sediments" may have contributed to the overall large change in polygon labels and the insignificant change in accuracy between the draft and contractor maps. This difficulty is especially evident in Phase I where the editing may have been skewed to include any possible areas of "SRV" habitat, thereby increasing the number of polygons changed to "SRV" in the contractor map and also increasing the number of "Unconsolidated Sediments" reference data points mapped as "SRV". However, since the "SRV/Unconsolidated Sediments" changed polygon group is not significantly larger than any other group of changed polygons, this particular habitat confusion is most likely not primary cause of the large change in polygons without an improvement in the accuracy during the editing process in Phase I of this project. Therefore, the null hypothesis for hypothesis one cannot be rejected.

Hypothesis 2: Accuracy Site Collection

In this study, it was found that the smallest transition zone has the highest percentage of changed polygons. Transition zone three, or any polygons within 0.1 meters horizontal distance of 0.5 meters depth, is only 9,822 hectares in size, but is made up of 99% changed polygons. Of the three transition zones, this zone also contains the largest area of changed polygons with 9,708 hectares. However, since the transition zone is so small, the large amount of changed polygons still only accounts for 13% of the total area of changed polygons. Regardless, the large number of changed polygons in transition zone three shows that the majority of editing was completed in the shallow waters around what may have been the shallow edges of seagrass beds.

This small zone also contains 159 accuracy sites, or 20% of the total number of accuracy sites. Since the percentage of the accuracy sites in this transition zone is larger than the percentage of the total changed area in this zone, it can be assumed that the transition zone is well sampled by the accuracy sites. Similarly, each of the deep water transition zones has a higher number of accuracy sites when compared to the amount of area in those zones. This analysis demonstrates that the researchers did not avoid transition zones when collecting accuracy sites, since they actually over-sampled the transition zones. Consequently, the null hypothesis for hypothesis two cannot be rejected as a result of these findings, since the test was to determine if the transition areas were under-sampled.

Hypothesis 3: Polygon Size

For both Phase I and Phase II, the average polygon size is always smaller for the original draft map, dissolved or not, than for the contractor map. However, for Phase II, this change is much less pronounced. In Phase I, the average polygon size doubled from the dissolved draft map to the contractor map, increasing from 4.289 hectares to 8.063 hectares (an 88% increase), while in Phase II the average polygon size only increased from 9.548 hectares to 10.589 hectares (an 11% increase). These averages might be skewed slightly higher than would be representative of the areas, since the distribution of the polygons is such that these areas have many smaller polygons and one or two larger ones. However, the distribution of polygon sizes in both the dissolved draft map and the contractor map appear similar in both Phase I and II (Appendix C). For Phase I, the skewed distribution of polygon sizes does not change the observation that the average size of the polygons increased during the editing process, but it does mean that the averages in Phase II may in actuality be even closer.

More importantly, the average sizes of the polygons that were changed between the draft and contractor maps are smaller than the average draft polygon size for both phases. In Phase I the average changed polygon size is 0.736 hectares, less than one-quarter the average draft polygon size. In Phase II, there was a similar result. The average changed polygon size is 2.252 hectares in Phase II, which is still smaller than the average draft polygon size. The smaller averages of the changed polygon sizes demonstrate that most of the editing was completed on the smaller polygons of the image. This result shows that many smaller polygons might have been assimilated into larger polygons during the editing process. However, many of the small polygons may have been wrongly absorbed into larger surrounding polygons if the researchers felt that the habitat was too small to be of concern for the final map, thereby creating a number of wrongly changed polygons.

A closer investigation of the editing process completed in Phase I shows that of all of the polygons with coinciding accuracy sites, the ones that were edited from wrong to right were generally larger than any of the polygons edited from right to wrong or wrong to wrong. However, all three of the changed polygon categories still average smaller in size than the average accuracy site size. This smaller average illustrates again that the changed polygons are typically smaller than the rest of the study area polygons. In addition, the different averages indicate that the larger polygons are easier to edit into the correct category.

Since the majority of the editing was completed on the smaller polygons, and the smaller the polygon the more likely it would not be edited into the right category, segmenting the image into very small polygons actually reduced the effectiveness of the mapping. As noted previously, this may be due to the fact that shadows, glint, or a number of other environmental factors, may split homogeneous habitats into two different polygons when the segmentation scale parameter is set too small (Schiewe, 2002). These two polygons of the same habitat could then be classified differently because of their slight difference in spectral reflectance.

The other issue with a small segmentation scale parameter has to do with the fact that during the process of defining the classification decision tree, the training points collected may not encompass all of the variation seen in a specific habitat class (Schiewe, 2002). With a larger segmentation parameter, the reflectance values of the pixels within a larger polygon are averaged and then that average is used to classify each polygon. In smaller polygons there is less information to be averaged since there are far fewer pixels. Therefore, with a smaller segmentation parameter there is more variation in each habitat class because there are a wider variety of polygons and less averaging of information. The wider variety of polygons can make capturing a habitat's natural variation much harder and can lead to the misclassification of many of these small polygons. The fact that as many of these small polygons in Phase I were changed from wrong to right as right to wrong during the editing process, further shows that the smaller polygons are much harder to correctly classify and edit. Therefore, the small initial segmentation scale parameter may be the culprit behind the negligible accuracy improvement in Phase I. Therefore, the third null hypothesis is the only rejected null hypothesis.

Conclusions

As determined by this research, the leading cause behind the large change in polygon labels between the draft and contractor maps, without an improvement in accuracy, was the choice in the segmentation scale parameter used when the researchers were preparing the original images for Phase I of this project. The Phase I segmentation size was too small to effectively map the benthic habitat study area. This assumption is reinforced by the fact that in the second phase of this project, the initial segmentation size was much larger than in Phase I, and the accuracy of the benthic habitat map vastly increased with far less editing.

Phase II of this project used larger initial polygons to create a benthic habitat map and this map was considerably easier to edit than the Phase I map. Although two separate levels of segmentation were implemented in Phase II, the time it took to create and edit the draft map was shorter than the time it took to create and edit the draft map in Phase I. The Phase II contractor map also resulted in a higher accuracy than the contractor map in Phase I. Given that it is quicker and more accurate, the Phase II method for classification, including the two levels of segmentation, is the more effective of the two mapping methods used for this research.

Therefore, the null hypothesis that can be conclusively rejected is null hypothesis three, showing that it is possible that the small initial segmentation scale parameter used to create the benthic habitat maps in Phase I may have caused the large number of changed polygons during editing without an improvement in accuracy. Null hypotheses one and two cannot be rejected and hypothesis one especially needs further investigation. In the exploration of hypothesis one, the confusion between "SRV" and "Unconsolidated

Sediments" seemed to be the most prevalent confusion between habitats, however no one specific group of changed habitats was statistically larger than any other group, indicating that the null hypothesis cannot be rejected and no particular pair of habitats caused the large change in polygons.

Although segmentation size appears to be the major contributor to the discrepancy between significant polygon editing and no increase in map accuracy, a smaller contributor was the confusion between the "SRV" and the "Unconsolidated Sediments" labels that was observed in both phases of the project. While this confusion may be an issue of location of habitats, more research should be initiated which focuses on how to better differentiate these habitats on aerial imagery. One dataset that may help in determining the difference between these habitats is low tide near-infrared (NIR) images. If the "SRV" habitats partially float on top of the water at low tide, the NIR image may more effectively capture the difference between healthy growing "SRV" and "Unconsolidated Sediments" (Macleod and Congalton, 1998).

In an effort to improve their own benthic habitat map, the researchers did complete a raster map of percent vegetative cover for the "SRV" areas. This method did prove slightly better at separating and identifying "Continuous SRV", "Patchy SRV", and "Unconsolidated Sediments", but it was not used in the creation of the benthic habitat maps in Phase I (Green et al., 2007). In the future, this method, as well as many other possible data sources, should be explored in an effort to help minimize the confusion caused by "SRV" and "Unconsolidated Sediments".

LITERATURE CITED

Anderson, J. R., E. E. Hardy, J. T. Roach, and R. E. Witmer. 1976. A Land Use and Land Cover Classification System for Use with Remote Sensor Data. Geological Survey Professional Paper 964. United State Geological Survey. Washington, DC. 28 pp.

Baatz, M., U. Benz, S. Dehghani, M. Heymen, A. Holtje, P. Hofmann, I. Ligenfelder, M. Mimler, M. Sohlbach, M. Weber, and G. Willhauck. 2001. *eCognition User Guide*. Munich: Definiens Imaging GmbH. 310 pp.

Blaschke, T. and J. Strobl. 2001. What's Wrong with Pixels? Some Recent Developments Interfacing Remote Sensing and GIS. *GIS*. Heidelburg: Huthig GmbH & Co. 6:12-17.

Bostrom, C., E. L. Jackson, C. A. Simenstad. 2006. Seagrass landscapes and their effects on associated fauna: a review. *Estuarine, Coastal and Shelf Science*. 68:383-403.

Chuvieco, E. and R. G. Congalton. 1988. Using Cluster Analysis to improve the Selection of Training Statistics in Classifying Remotely Sensed Data. *Photogrammetric Engineering and Remote Sensing*. 54(9):1275-1281.

Congalton R.G. 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*. 37:35-46.

Congalton, R.G. and K. Green. 1992. The ABCs of GIS: An introduction to geographic information systems. *Journal of Forestry*. 90/11:13-20.

Congalton, R.G. and K. Green. 2009. *Assessing the accuracy of remotely sensed data: principles and practices, Second Edition*. CRC Press, Boca Raton, Fl. 208 pp.

Congalton, R. G., R. G. Oderwald, and R. A. Mead. 1983. Assessing Landsat classification accuracy using discrete multivariate statistical techniques. *Photogrammetric Engineering and Remote Sensing*. 49(12):1671-1678.

Cowardin, L. M., V. Carter, F. C. Golet, and E. T. LaRoe. 1979. Classification of Wetlands and Deepwater Habitats of the United States. Office of Biological Services. US Fish and Wildlife Service. FWS/OBS-79/31. US Department of the Interior, Washington, DC. 103 pp.

De'ath, G. and K. E. Fabricius. 2000. Classification and Regression Trees: a powerful yet simple technique for ecological data analysis. *Ecology*. 81(11):3178-3192.

Diaz, R. J., M. Solan, and R. M. Valente. 2004. A review of approaches for classifying benthic habitats and evaluating habitat quality. *Journal of Environmental Management*. 73:165-181.

Divins, D.L., and D. Metzger. NGDC Coastal Relief Model. June 30, 2008. *http://www.ngdc.noaa.gov/mgg/coastal/coastal.html*.

Dobson, J. E., E. A. Bright, R. L. Ferguson, D. W. Field, L. L. Wood, K. D. Haddad, H. Iredale III, J. R. Jensen, V. V. Klemas, R. J. Orth, and J. P. Thomas. 1995. *NOAA Coastal Change Analysis Program (C-CAP): Guidance for Regional Implementation*. NOAA Technical Report NMFS 123. US Department of Commerce. Seattle, WA.

Ferguson, R. and K. Korfmacher. 1997. Remote sensing and GIS analysis of seagrass meadows in North Carolina, USA. *Aquatic Botany*. 58:241-258.

Florida Department of Transportation, State Topographic Bureau, and Thematic Mapping Section. 1985. *Florida Land Use, Cover and Forms Classification System*. Procedure No. 550-010-001-a, Second Edition. Florida Department of Transportation.

Florida Natural Areas Inventory and Florida Department of Natural Resources. 1990. *Guide to the Natural Communities of Florida*. Florida Natural Areas Inventory.

Friedl, M. A. and C. E. Brodley. 1997. Decision Tree classification of Land Cover from Remotely Sensed Data. *Remote Sens. Environ*. 61:388-409.

Friedlander, A. M., E. K. Brown, and M. E. Monaco. 2007. Coupling Ecology and GIS to Evaluate Efficacy of Marine Protected Areas in Hawaii. *Ecological Applications*. 17(3):715-730.

Green, K. and C. Lopez. 2007. Using object-oriented classification of ADS40 data to map the benthic habitats of the state of Texas. *Photogrammetric Engineering & Remote Sensing*. 861-865.

Green, K., C. Lopez, and H. Rempel. 2007. *Coastal Bend of Texas Benthic Habitat Mapping Phase 1 Final Report*. Unpublished Manuscript. 54pp.

Green, K., C. Lopez, and H. Rempel. 2008. *Lower Laguna Benthic Mapping Accuracy Assessment Analysis*. Unpublished Manuscript. 6pp.

Handley, L., Altsman, D., and DeMay, R., eds. 2007. *Seagrass Status and Trends in the Northern Gulf of Mexico: 1940-2002*. U.S. Geological Survey Scientific Investigations Report 2006-5287 and U.S. Environmental Protection Agency 855-R-04-003. 267 pp.

Hansen, M. R. Dubayah, and R. DeFries. 1996. Classification trees: an alternative to traditional land cover classifiers. *Int. J. Remote Sensing*. 17(5):1075-1081.

Jensen, J. R. 2005. *Introductory Digital Image Processing: A Remote Sensing Perspective, Third Edition*. Pearson Prentice Hall, Upper Saddle River, NJ. 526 pg.

Koch, E. 2001. Beyond Light: Physical, Geological, and Geochemical Parameters as Possible Submersed Aquatic Vegetation Habitat Requirements. *Estuaries*. 24(1):1-17.

Lee, K. S., S. R. Park, Y. K. Kim. 2007. Effects of irradiance, temperature, and nutrients on growth dynamics of seagrasses: a review. *Journal of Experimental Marine Biology and Ecology*. 350:144-175.

Macleod, R. D., and R. G. Congalton. 1998. A Quantitative Comparison of Change-Detection Algorithms for Monitoring Eelgrass from Remotely Sensed Data. *Photogrammetric Engineering & Remote Sensing*. 64(3):207-216.

Madley, K.A., B. Sargent, and F.J. Sargent. 2002. Development of a System for Classification of Habitats in Estuarine and Marine Environments (SCHEME) for Florida. Unpublished report to the U.S. Environmental Protection Agency, Gulf of Mexico Program (Grant Assistance Agreement MX-97408100). Florida Marine Research Institute, Florida Fish and Wildlife Conservation Commission, St. Petersburg. 43pp.

Mumby. P. J., E. P. Green, A. J. Edwards, and C. D. Clark. 1997. Coral reef habitat mapping: how much detail can remote sensing provide?. *Marine Biology*. 130:193-202.

Mumby, P. and A. R. Harborne. 1999. Development of a systematic classification scheme of marine habitats to facilitate regional management and mapping of Caribbean coral reefs. *Biological Conservation*. 88:155-163.

National Oceanic and Atmospheric Administration (NOAA) Coastal Services Center. 2008. *Benthic Habitat Mapping*. http://www.csc.noaa.gov/benthic/.

Onuf, C. P. 1996. Seagrass responses to long-term light reduction by brown tide in upper Laguna Madre, Texas: distribution and biomass patterns. *Marine Ecology Progress Series*. 138:219-231.

Plourde, L. and R. G. Congalton. 2003. Sampling Method and Sample Placement: How Do They Affect the Accuracy of Remotely Sensed Maps?. *Photogrammetric Engineering and Remote Sensing*. 69(3):289-297.

Rahman, A. F., J. A. Gamon, D. A. Sims, and M. Schmidts. 2001. Optimum pixel size for hyperspectral studies of ecosystem function in southern California chaparral and grassland. *Remote Sensing of the Environment*. 84:192-207.

Schiewe, J. 2002. Segmentation of high-resolution remotely sensed data – concepts, applications, and problems. In the ISPRS Symposium on "Geospatial Theory, Processing and Applications". Ottawa.

Story, M. and R. G. Congalton. 1986. Accuracy Assessment: A User's Perspective. *Photogrammetric Engineering and Remote Sensing*. 52(3):397-399.

Warner, T. A., J. Y. Lee, and J. B. McGraw, 1998. Delineation and identification of individual trees in the Eastern Deciduous Forest. *International Forum on Automated Interpretation of High Spatial Resolution Imagery for Forestry*. February 10-12, 1998, Victoria B.C.

Zajac, R. N., R. S. Lewis, L. J. Poppe, D. C. Twitchell, J. Vozarik, and M. L. DiGiacomo-Cohen. 2003. Responses of infaunal populations to benthoscape structure and the potential importance of transition zones. *Limnol. Oceanogr*. 48(2):829-842.

APPENDICES

APPENDIX A

THE FLORIDA SYSTEM FOR CLASSIFICATION OF HABITATS IN ESTURARINE AND MARINE ENVIRONMENTS (SCHEME) (Madley et al., 2002)

Classification categories are structured as follows:
X Class
  XX Subclass 1
    XXX Subclass 2
      XXXX Subclass 3
        XXXXX Subclass 4

Habitats:

1. Unconsolidated Sediments (zero to less than 10 percent colonization)
Unconsolidated sediments with less than 10 percent colonization by SAV or corals.

11. Mud (i.e. silts and clays) [<0.0625mm grain sizes comprise greater than 50% of sediment]
Sediments most often found in depositional environments that are protected from wind and wave energy.

12. Sand [0.0625-2mm grain sizes comprise greater than 50% of sediment]
Sediments usually found in areas exposed to wind and wave energy that causes silts and clays to be removed.

121. Very fine sand [0.0625-0.125mm grain sizes comprise greater than 50% of sediment]

122. Fine sand [0.125-.25mm grain sizes comprise greater than 50% of sediment]

123. Medium sand [0.25-.5mm grain sizes comprise greater than 50% of sediment]

124. Coarse sand [0.5-1mm grain sizes comprise greater than 50% of sediment]

125. Very coarse sand [1-2mm grain sizes comprise greater than 50% of sediment]

13. Mixed Fine

Mixture of sand and mud, possibly with sparse grains of larger size categories such as granules or pebbles (no one substrate type presence is greater than 50%)

14. Mixed Coarse

Granules, pebbles, cobbles are the possible components that comprise over 50% of the sediment (no one substrate type presence is greater than 50%)

141. Shell hash – substrate covered with a mixture of shell material from granules (2-4mm grain sizes) to whole shells.

15. Granule [2-4mm grain sizes comprise greater than 50% of sediment]

16. Pebble [4-64mm grain sizes comprise greater than 50% of sediment]

17. Cobble [64-256mm grain sizes comprise greater than 50% of sediment]

Note: Rocks larger than 256mm (=10 in) in diameter are classified as bedrock in the consolidated bottom category.

18. Detrital floor

Detrital material (e.g. seagrass, algae, leaf litter, etc.) that builds up in intertidal and shallow waters, often along windward shorelines. This semi-permanent feature creates an organic mud buildup under the newly deposited detrital material. More permanent feature than the Drift Wrack general modifier.

2. Submersed Aquatic Vegetation (SAV)

Any combination of SAV (i.e. seagrasses, oligohaline grasses, attached macroalgae and drift macroalgae) that covers 10 to 100 percent of a substrate. If reef or hardbottom is more abundant than the SRV the polygon should be recorded as Reef/Hardbottom Class and SRV should be noted with Modifiers.

21. Submersed Rooted Vascular Plants (SRV) (i.e. seagrasses and oligohaline grasses)

Habitat with 10 percent or more cover of SRV.

211. Continuous SRV

This includes continuous beds of any shoot density (i.e. sparse continuous, dense continuous or any combination). These areas appear as continuous seagrass signatures; however, small (less than 0.5 acres) bare sediment areas may be observed as infrequent features within the area.

2111. Dense patches of SRV in a matrix of continuous, sparse SRV

Continuous coverage of sparse SRV in which dense patches of SRV are clearly observed interspersed within the area. This pattern is often the result of effects from the sediment or underlying bedrock characteristics.

212. Discontinuous SRV

Areas of SRV with breaks in coverage that result in isolated patches of SRV, usually in unconsolidated bottom but also exist in hard bottom areas. If the hardbottom is more abundant than the SRV the polygon should be recorded as Reef/Hardbottom Class and SRV can be noted with Modifiers. Generally, these grass features appear as semi-round patches or elongated strands separated by bare sediment.

22. Macroalgae

221. Attached Macroalgae

Habitat with 10 percent or more cover of mixed or monospecific macroalgae attached to the substrate with holdfasts, rhizomes, or other morphological feature.

2211. Continuous attached macroalgae

This includes continuous beds of any density (i.e. sparse continuous, dense continuous or any combination). These areas appear as continuous attached macroalgae or SRV signatures. Often macroalgae can't be interpreted from the imagery without field verification to detect the difference from SRV. Small (less than 0.5 acres) bare sediment areas may be observed as infrequent features within the area.

22111. Dense patches of attached macroalgae in a matrix of continuous, sparse macroalgae

Continuous coverage of sparse attached macroalgae in which dense patches of attached macroalgae are clearly observed interspersed within the area. This pattern is often the result of effects from the sediment or underlying bedrock characteristics.

2212. Discontinuous attached macroalgae

Areas of attached macroalgae with breaks in coverage that result in isolated patches, usually in unconsolidated bottom but also exist in hard bottom areas.

222. Drift Macroalgae

Habitat with 10 percent or more cover of mixed or monospecific macroalgae that is not attached to the substrate. Drift algae may move constantly with wind or wave forces or may be observed in one location for long periods of times (possibly months) because of lack of energy forces or due to becoming entangled on substrate features.

2221. Continuous drift macroalgae

This includes continuous beds of any density (i.e. sparse continuous, dense continuous or any combination). These areas appear as continuous attached macroalgae or SRV signatures. Often macroalgae can't be interpreted from the imagery without field verification to detect the difference from SRV. Small (less than 0.5 acres) bare sediment areas may be observed as infrequent features within the area.

2222. Discontinuous drift macroalgae

Areas of attached macroalgae with breaks in coverage that result in isolated patches, usually in unconsolidated bottom but also exist in hard bottom areas.

3. Reef/Hardbottom

Hardened substrate of unspecified relief formed by the deposition of calcium carbonate by reef building corals and other organisms or exposed bedrock, possibly with various degrees of concealment from attached plant and animal colonization. Unconsolidated bottom and SAV may occur within these habitats, although in less abundance than the reef/hardbottom.

31. Coral Reef

Hardened substrate formed by reef building corals. May be live coral or relict reefs. Often bedrock is the base for these reefs but the presence of coral or remnant coral on the surface is reason to categorize the dominant habitat as coral reef.

311. Platform Reef (also bank reef)

Hardened substrate formed by reef building corals that exist in a quasi-continuous structure along a shelf edge or similar dropoff removed from any coastline. These are typically elongate structures and may be referred to as bank reefs. The following Subclass categories may be present in various combinations within a platform reef.

3111. Linear Reef

Linear, contiguous coral formations. Reef crest, fore reef, and back reef zones could be mapped as Linear Reef. Most often has associated spur and groove and reef rubble habitats.

31111. Reef Terrace (high profile)

Contiguous reef with high complexity and high relief (>2m).

31112. Remnant (low profile)

Reefs of relief less than 2m that lack distinctive spur and groove characteristics. These reefs consist of coral and hard bottom features; often support soft corals, sponges, seagrass; and are usually found growing parallel to the reef tract, though they may form transverse features that grow perpendicular to the reef tract.

3112. Spur and Groove

Distinct coral bands separated by sand or uncolonized hardbottom grooves. This habitat type usually occurs in the fore reef zone.

31121. High Relief Spur and Groove

Distinct coral bands separated by sand or uncolonized hardbottom grooves. The coral bands have 1.5- 4m relief.

31122. Low Relief Spur and Groove

Distinct coral bands oriented perpendicular to the shore or bank and separated by sand or uncolonized hardbottom grooves. The coral bands have <1.5m relief.

3113. Reef Rubble

Dead, unstable coral rubble that often occurs landward of platform reefs.

312. Patch Reefs

Irregularly shaped reef communities. They may range in size from tens to thousands of square meters. Patches are separated from each other by uncolonized hardbottom, sand, or colonized substrate with submersed aquatic vegetation (SAV), macroalgae, gorgonians or sponges. Most often the patches are surrounded by a halo of bare substrate created by foraging, obligate reef inhabitants.

3121. Individual Patch Reef

Isolated, single reef (larger than the minimum mapping unit of the project) without associated halo area. These individual reefs may have an associated halo, however if large enough (i.e. greater than the minimum mapping unit) to be delineated the halo will be mapped as its own subclass.

3124. Aggregated Patch Reefs (includes Halo areas if present)

Clustered patch reefs, usually too small (less than the minimum mapping unit) or too close together to map individually or where halos coalesce.

3125. Pinnacles

High complexity patch reefs that have high relief (up to 15m) from the sea floor. These structures may occur in clusters and are typically surrounded by large sand plains.

313. Patchy Coral and/or Rock in Unconsolidated Bottom

Areas of primarily sand, submerged aquatic vegetation (SAV), or low relief rock covered with a sand veneer. Often adjacent to spur and groove habitats, these areas contain small, individual corals or rocks that are distinctive yet a very low percentage of the total cover (and certainly smaller than the minimum mapping unit).

32. Mollusk Reefs

Concentrations of sessile mollusks that attach to hard substrate and with the correct conditions will proliferate allowing the reef to grow. In Florida, these areas are most common in estuarine areas and are not known to occur in water deeper than 40 feet.

321. Bivalve Reefs (i.e. oyster reefs)

Mollusk reefs dominated by oysters; at times partially exposed during low tide.

322. Gastropod Reefs (i.e. Vermetid reefs)

Mollusk reefs created by a worm-like mollusk of the genus Petaloconchus. In Florida, these reefs are only known to be found in shallow waters seaward of the outer islands in the Ten Thousand Islands area of southwest Florida.

33. Annelid Reefs (i.e. Sabellariid reefs)

Structures formed from colonies of Sabellariid worm tubes. Commonly found in the tidal zone on the east coast of Florida, these structures are mostly formed on hard substrates and may be exposed at low tide. Storm events can break the sand structures thus changing the extent of the colony at the time of mapping. The reefs also expand as worm larvae settle on the mounds and build additional tubes.

34. Hardbottom

Hard substrate composed of exposed bedrock or created through syndepositional cementation of sediment.

341. Bedrock

Exposed bedrock and/or rocky outcrops with low to high relief and high complexity.

342. Pavement (i.e. low relief hardbottom)

Flat, low relief, mostly solid rock substrate.

4. Tidal Marsh (i.e. salt marsh, coastal marsh)

Communities of emerged halophytic vegetation along low-wave energy intertidal areas and river mouths. These areas are dominated by grasses, rushes and sedges (i.e cordgrass, needlerush, and sawgrass).

41. Salt pan

Exposed or water-filled depressions in a tidal marsh area. Often covered by layers of blue-green algae but possibly bare sediment only. Glassworts or saltworts may be present. Sand barrens most often exist in high marsh areas; conversely mud barrens may occur in the intertidal zone as water retention pools during low tide.

42. Salt marsh algae

Mud flats dominated by a mixture of benthic microalgae, phytoplankton, and macroalgae.

5. Tidal Swamp (i.e. mangrove, mangrove forest)

Dense, low forests primarily located along coastal areas. Various tidal marsh grasses and shrubs may be associated but these communities are dominated by a mix of red, black and white mangroves.

6. Land

Mainland, islands, causeways and other land normally above the high tide line. Depending on the mapping project the line delineating the water/land interface may be formed anywhere between the extreme low and extreme high tide marks.

7. Unknown benthic habitat (i.e. not lending to interpretation because of water quality, depth, or lack of field investigation)

71. Turbid plume

Area of dark colored water often associated with river mouths, bays and coastlines. The often brown water results from vegetation tannins leached into the rivers and/or organic particles carried seaward from the river water. Plumes varying in color from white to emerald green are also observed in areas with fine carbonate sediments (e.g. Florida keys). These plumes will often prevent mapping of benthic habitats from photography.

72. Phytoplankton bloom

Area of water that contains abnormally high concentrations of phytoplankton. These blooms can often be seen in aerial photography and will prevent mapping of benthic habitats by decreasing water clarity. The classes listed below should be used only if a distinct color is associated with the bloom when mapped.

721. Green bloom

722. Red bloom

723. Black bloom

<u>**General Modifiers**</u> – Modifier labels (e.g. A, B, C…) will be used to indicate more specific information about map categories. For instance, 123-E would represent a natural, submersed tidal channel with medium sand sediment type. Also, an area of flat, low relief hardbottom with dominant cover types of octocorals and sponges would be labeled 342-QR. Another example, 34-A22 would represent an artificial reef consisting of concrete culverts.

A. Artificial

- **1. Tires**
- **2. Concrete materials of opportunity**
  - **21. Concrete blocks**
  - **22. Culverts**
  - **23. Riprap**
- **3. Designed materials**
  - **31. Reef balls**
  - **32. PVC structures**
- **4. Vessels, automobiles, planes, military ordnance (whole or portions)**
- **5. Steel structures (i.e. oil rigs, lighthouses, etc.)**
- **6. Cables**
- **7. Pipelines**

B. Venetian canals – anthropogenic canals landward of shoreline

C. Streams – natural canals landward of shoreline

D. Island moats – deepwater canals wholly or partially surrounding islands

E. Submersed tidal canals – natural canals seaward of shoreline

F. Dredged/Excavation – anthropogenic canals seaward of shoreline or sediment borrow pits

G. Spoil/Fill – area of positive vertical relief created by placement of dredged sediments

H. Restoration – area of restoration activity (e.g. previous salt marsh planting area)

I. Seagrass – this modifier can be used when seagrass is present in areas of 10% or greater cover of coral or hardbottom

J. Drift seagrass – accumulation of seagrass that may be drifting in water column or lying on the bottom

K. Drift wrack – mix of various materials (e.g. seagrass, macroalgae, mangrove litter, etc.) that may be drifting in water column or lying on the bottom

L. Drift macroalgae – accumulation of macroalgae that may be drifting in water column or lying on the bottom

M. Mat algae – thin veneer of algae on substrate

N. Attached Macroalgae

1. **mixed browns**
2. **mixed reds**
3. **mixed greens and calcareous**

O. Urchin Front – congregation of urchins often dense enough to obscure the substrate in photography

P. Boat propeller scars

1. **Light scarring** – scarring in less than 5% of an SAV polygon
2. **Moderate scarring** – scarring in 5-20% of an SAV polygon
3. **Heavy scarring** – scarring in more than 20% of an SAV polygon

Q. Octocoral bed – soft coral species attached to substrate

R. Sponge bed – sponge species attached to substrate

S. Hard corals – hard coral species attached to substrate

T. Dead coral

U. Substrate ripples – area containing troughs and ridges of substrate as opposed to flat substrate. Height range listed in meters e.g. (0.5-1.0m)

V. Carbonate substrate

W. Siliciclastic substrate – pertaining to clastic noncarbonate rocks or sediments which are almost exclusively silicon-bearing, either as forms of quartz or as silicates

X. Sediment depth – depth of the surface substrate material. Range listed in meters e.g. (0-2m)

Y. Salinity (e.g. 32ppt)

Z. Tidal Status

1. **Subtidal** – never exposed to the air
2. **Intertidal** – exposed to the air even if only during the lowest spring tides

3. **Supratidal** – normally exposed to the air, only submersed during flood or storm events

AA. Biological habitat modifications

1. **Fish excavations** (e.g. tilefish, grouper, stingrays)
2. **Invertebrate bioturbation zones** (e.g. polychaetes, crabs, shrimps, etc.)

BB. SAV epiphytes – presence of algal or animal epiphytes are visible on the surface of SAV. A relative abundance can be characterized with the sub-modifiers listed:

1. **Light**
2. **Medium**
3. **Heavy**

<u>Taxonomic Modifiers</u> – This is a partial list. The intent is to build this list from expert input or compilation of published Florida species lists. These taxonomic modifiers would be used similar to the General Modifiers list except labels would be composed of the genus and species abbreviations composed of the first letter of the genus and the first three letters of the species. This coding mechanism allows for inclusion of new species and manageable changes as taxonomic name changes occur through peer-reviewed literature.

Examples: 211-Ttes would represent an area of dense, continuous turtle grass (*Thalassia testudinum*) whereas 211 would represent an area of dense, continuous SRV of unknown species types. Likewise, an area of coastal needlerush marsh would be labeled 4-Jroe. These modifiers would be used in order of decreasing percent cover. For instance, 4-Jroe/Salt would indicate a polygon of tidal marsh with a mixture of needlerush (*Juncus roemerianus*) and smooth cordgrass (*Spartina alterniflora*); needlerush is more prevalent in this polygon because the Jroe modifier occurs first in the sequence.

Seagrasses
Thalassia testudinum
Halodule wrightii
Syringodium filiforme
Halophila engelmanni
Halophila johnsonii
Halophila decipiens

Tidal Marsh Plants
Ruppia maritima
Vallisneria americana
Juncus roemerianus
Spartina alterniflora
Spartina patens
Salicornia virginica

Cladium mariscoides
Batis maritima
Distichlis spicata
Salsola kali

Tidal Swamp Plants
Avicennia germinans
Laguncularia racemosa
Rhizophora mangle
Conocarpus erecta

Macroalgae
Argardhiella spp.
Avrainvellea spp.
Batophora spp.
Bryopsis spp.
Calothrix spp.
Caulerpa spp.
Chondria spp.
Cladophora spp.
Dictyota spp.
Digenia spp.
Gracilaria spp.
Halimeda spp.
Laurencia spp.
Oscillatoria spp.
Penicillus spp.
Rhipocephalus spp.
Sargassum spp.

Corals
Acropora cervicornis
Acropora palmata
Acropora spp.
Agaricia spp.
Montastrea annularis
Oculina varicosa
Porites porites
Porites spp.
Lophelia prolifera
Enallopsammia profunda
Siderastrea spp.

APPENDIX B

THE DICHOTOMOUS KEY THAT PROVIDES THE RULES USED TO CLASSIFY HABITATS IN THE SEAGRASS MONITORING PROGRAM (Green et al., 2007)

List of Habitats

Value	Habitat
0	Unclassified
1	Unconsolidated Sediments
211	Continuous Submerged Rooted Vascular Vegetation (SRV)
212	Patchy SRV
2111	Dense Patches of SRV in matrix of continuous sparse SRV
221	Continuous Macroalgae
223	Patchy Macroalgae
3	Reef/Hardbottom
32	Mollusk Reef
321	Bivalve reefs (oyster)
34	Hardbottom
4	Tidal Marsh – Spartina
5	Tidal Swamp – Mangroves
6	Land
7	Unknown benthic habitat

All applicable SCHEME modifiers

The minimum mapping unit for this project is 0.01 hectares.

The Dichotomous Key

If habitat is falls within the "land" boundary as identified either by image classification or ancillary data then

If landcover consists of greater than or equal to 50% oysters, then **Bivalve Reef (321)**
Else if landcover is greater than or equal to 50% mangrove tree canopy, then **Tidal Swamp-Mangroves (5)**
Else if landcover is greater than or equal to 50% spartina, then **Tidal Marsh – Spartina (4)**
Else **Land (6)**

Else benthic habitat

If interpretation of benthic habitat is not possible because of water quality or water depth, then **Unknown Benthic Habitat (7)**

If Submerged Aquatic Vegetation (SAV) cover is greater than 10%, and reef/hardbottom cover is less than SAV cover then SAV (2)[1]

If greater than 10% of cover consists of Submerged Rooted Vascular Plants (SRV), then SRV (21).

If SRV cover is greater than or equal to 75% total coverage, then **Continuous SRV (211)**

If dense patches of SRV (greater than or equal to 75% coverage within patch) are interspersed within a matrix of sparse SRV (10-50% coverage), then **Dense Patches of SRV in matrix of Continuous Sparse SRV (2111)**

Else **Patchy (Discontinuous) SRV (212)**

Else if greater than 10% of cover consists of attached or drift macroalgae, then Macroalgae (22)

If Macroalgae cover is greater than or equal to 90% total coverage, then **Continuous Macroalgae (221)**
Else **Patchy (Discontinuous) Macroalgae (223)**

Else if SAV cover is greater than 10% and reef/hardbottom cover is greater than SAV cover, then **Reef/Hardbottom (3) with appropriate SAV modifiers.**

Else if SAV cover is less than 10%, and unconsolidated sediments make up more than 90% of substrate, then **Unconsolidated Sediments (1).**

Else if SAV cover is less than 10% and greater than 10% of substrate is reef/hardbottom, then Reef/Hardbottom (3)

If reef/hardbottom consists of greater than 50% sessile mollusks, then Mollusk Reef (32)

If mollusk reef consists of greater than 50% oysters, then **Bivalve Reef (321)**
Else **Mollusk Reef (32)**

Else if reef/hardbottom consists of greater than 50% exposed bedrock or hardbottom created by syndepositional cementation of sediment, then **Hardbottom (34)**
Else **Reef/Hardbottom (3)**

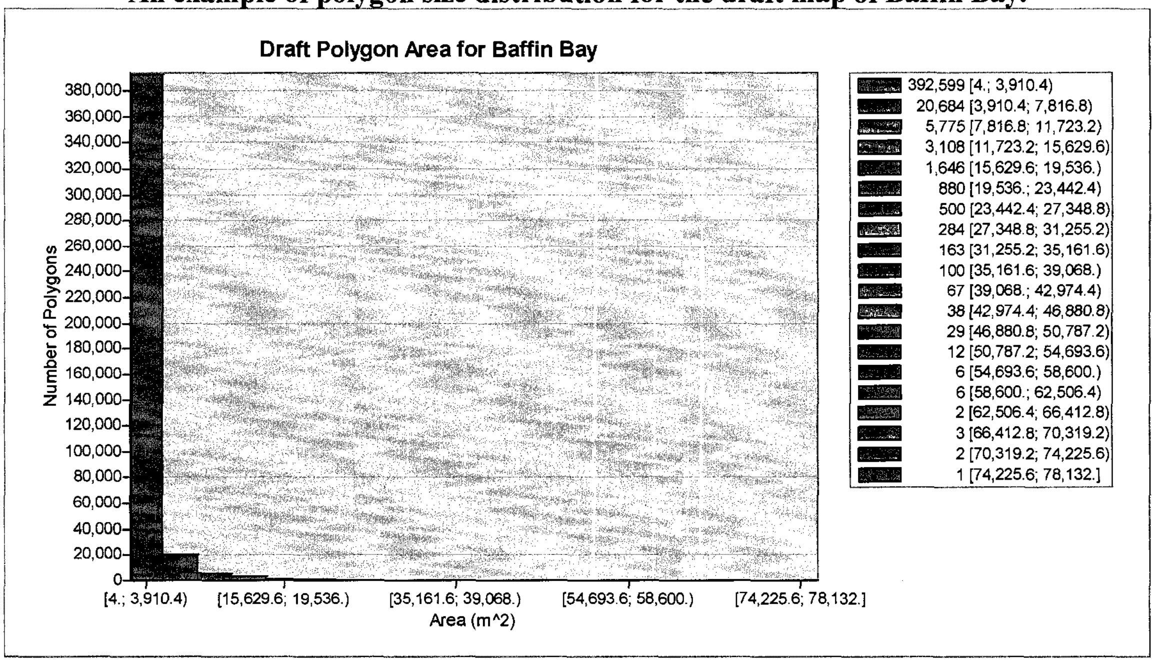
Else **Unknown Benthic Habitat (7)**

APPENDIX C

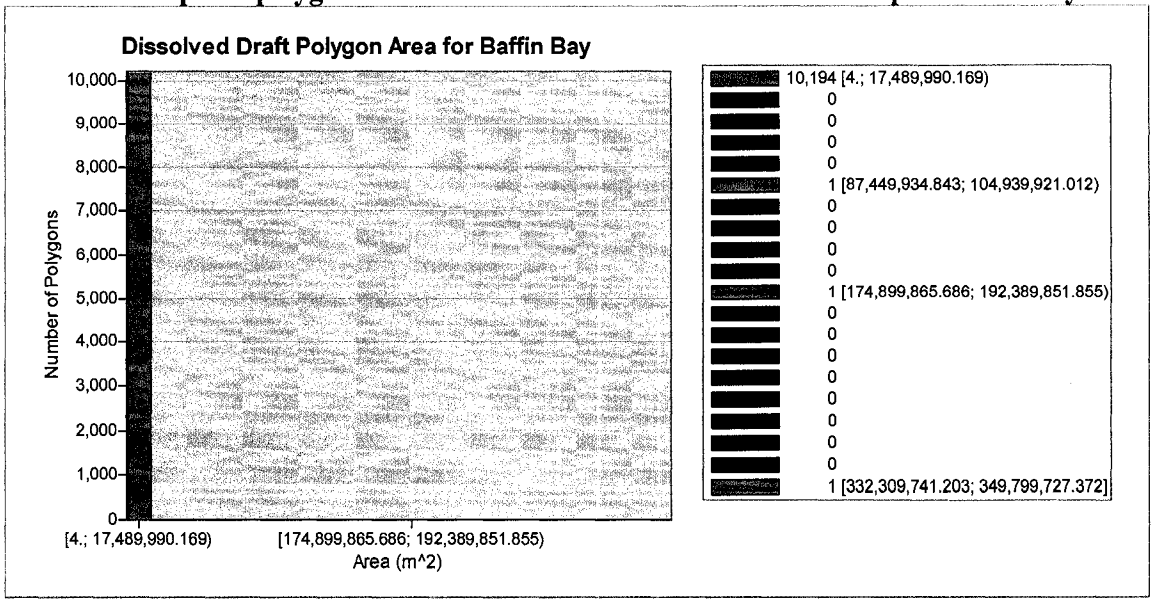
POLYGON SIZE DISRIBUTION FOR EACH MAP IN PHASES I AND II

Phase I

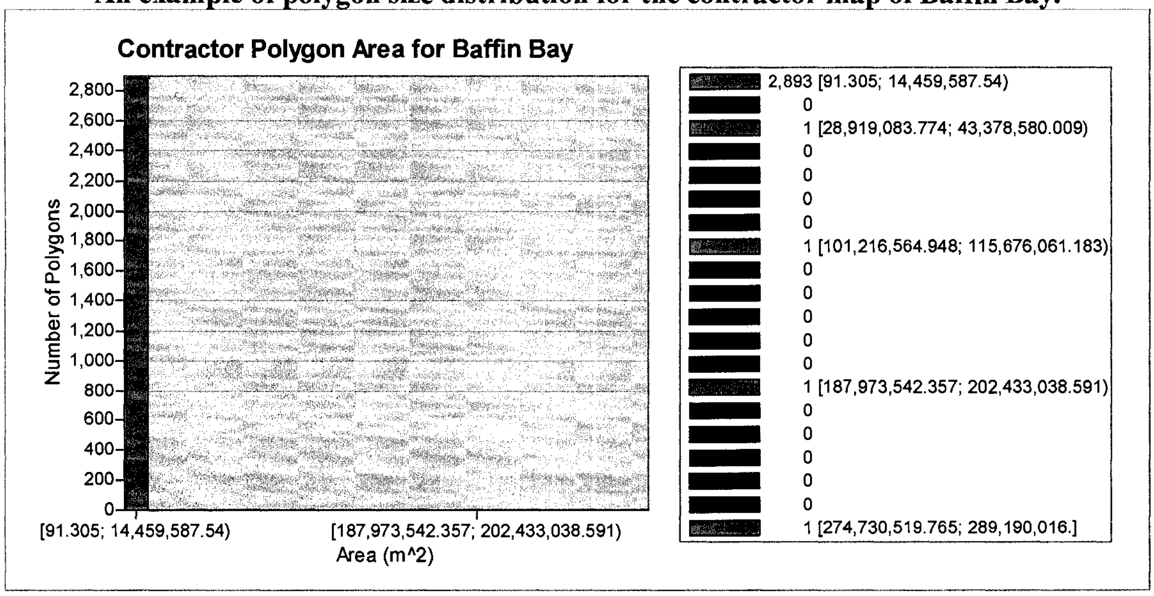
An example of polygon size distribution for the draft map of Baffin Bay.



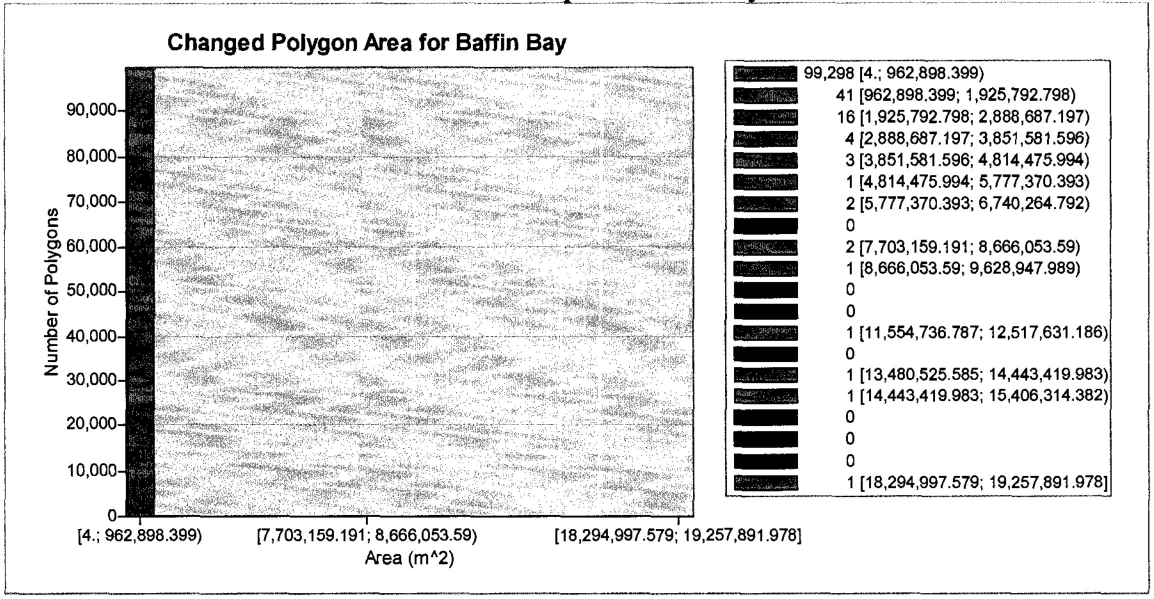
An example of polygon size distribution for the dissolved draft map of Baffin Bay.



An example of polygon size distribution for the contractor map of Baffin Bay.

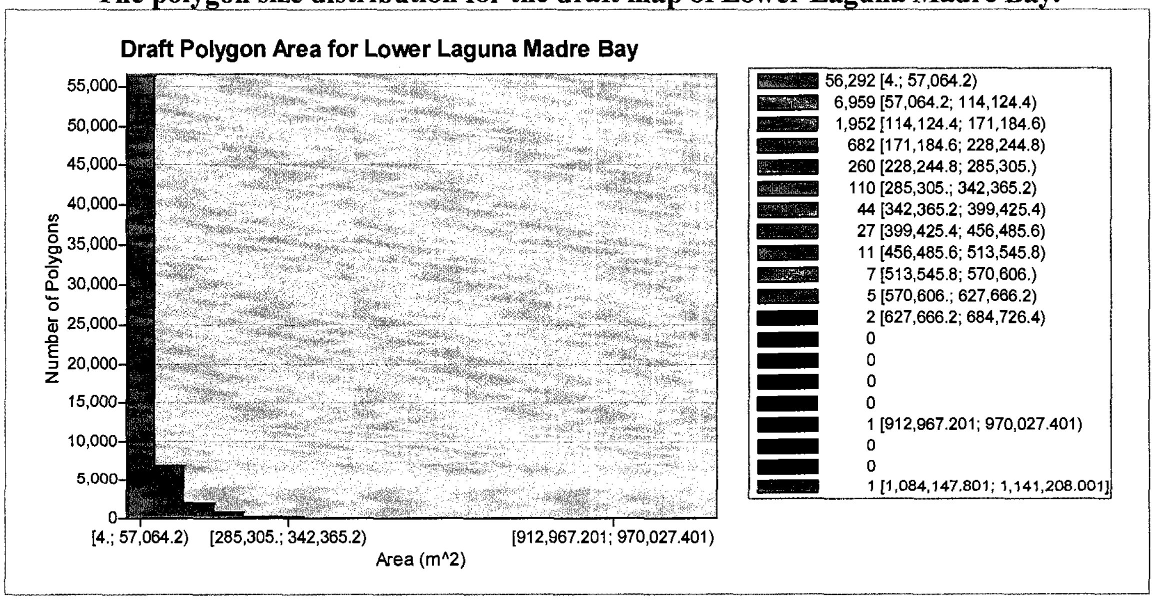


An example of polygon size distribution for the changed polygons between the draft and contractor maps of Baffin Bay.

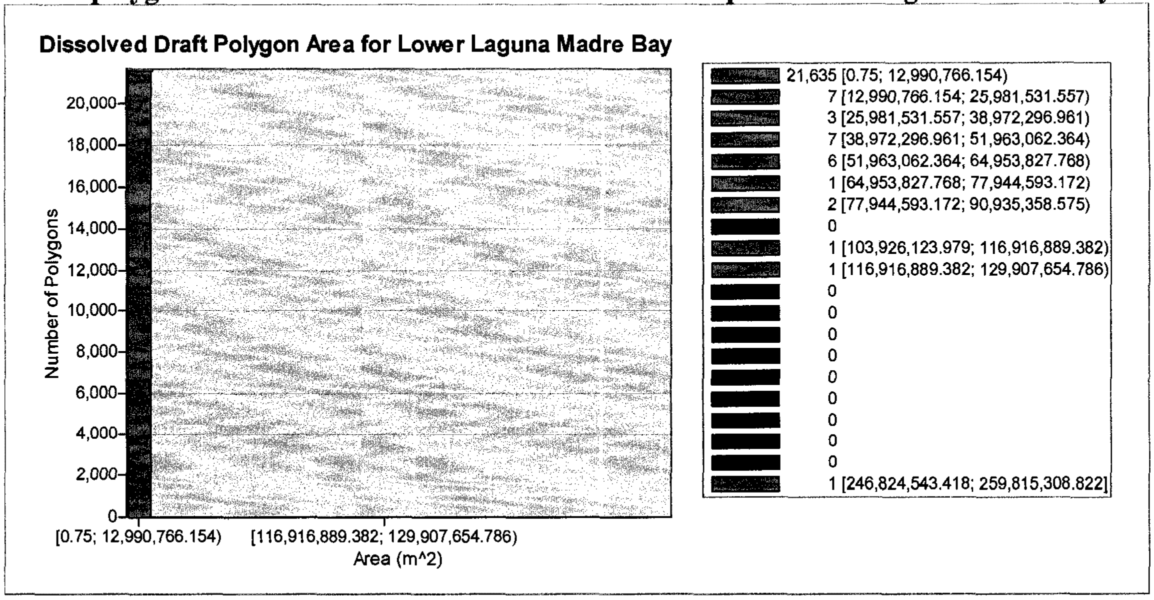


Phase II

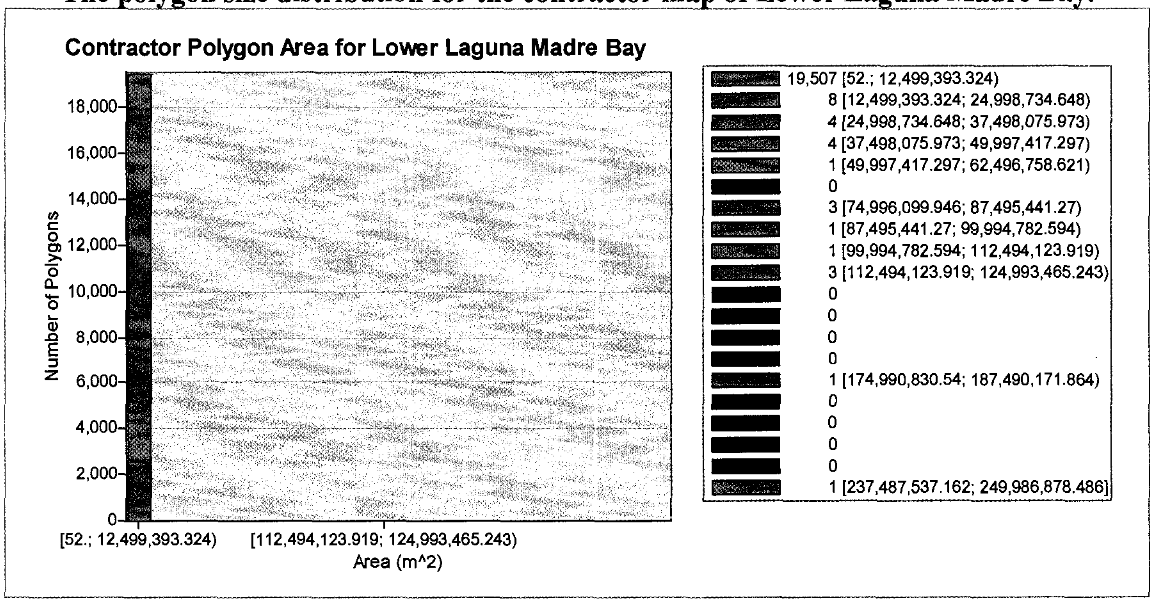
The polygon size distribution for the draft map of Lower Laguna Madre Bay.



The polygon size distribution for the dissolved draft map of Lower Laguna Madre Bay.



The polygon size distribution for the contractor map of Lower Laguna Madre Bay.



The polygon size distribution for the changed polygons between the draft and contractor maps of Lower Laguna Madre Bay.